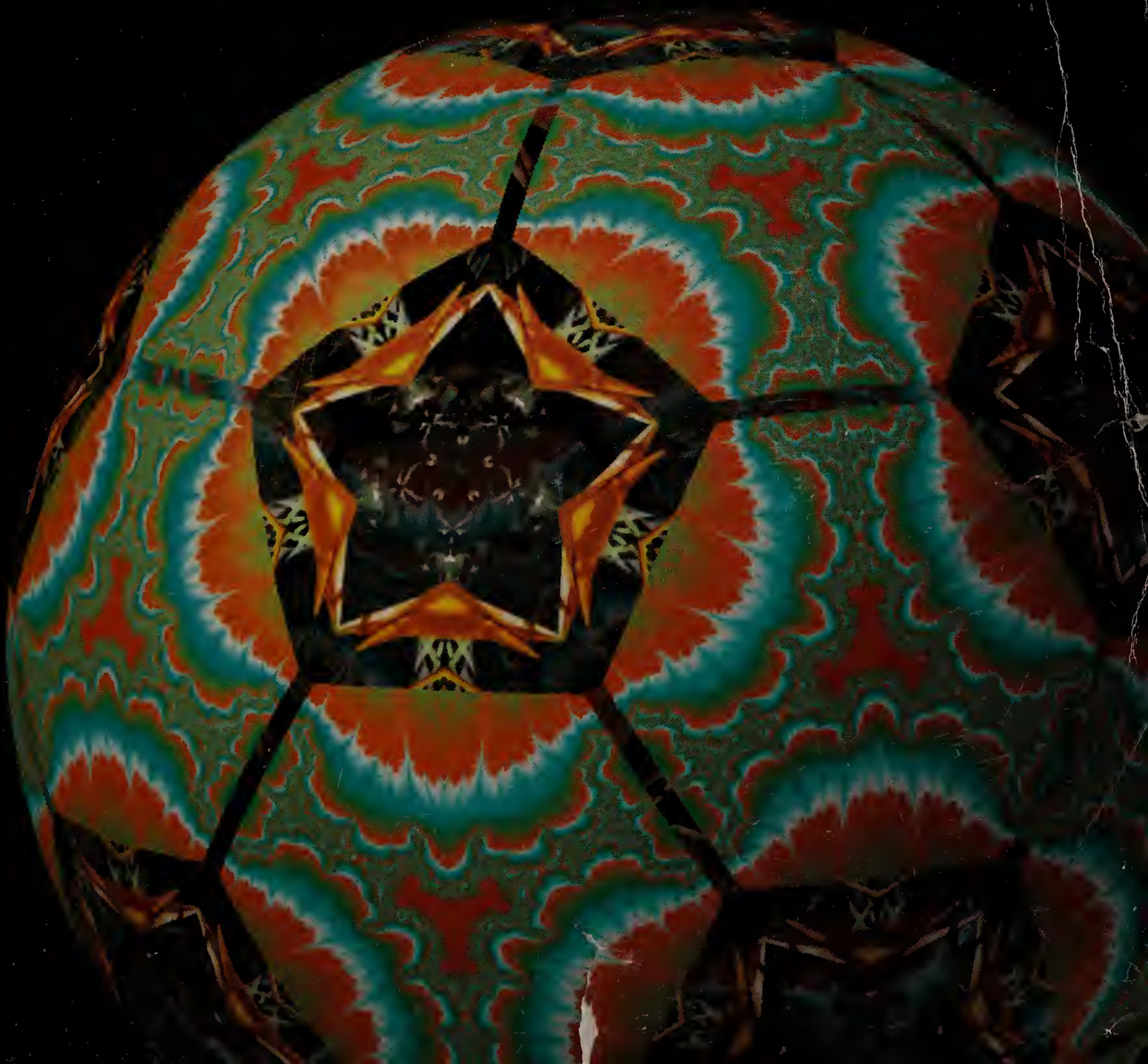
PHILLIPS ACADEMY
3 1867 00713 5077

rt of Problem Solving

Precalculus

Richard Rusczyk



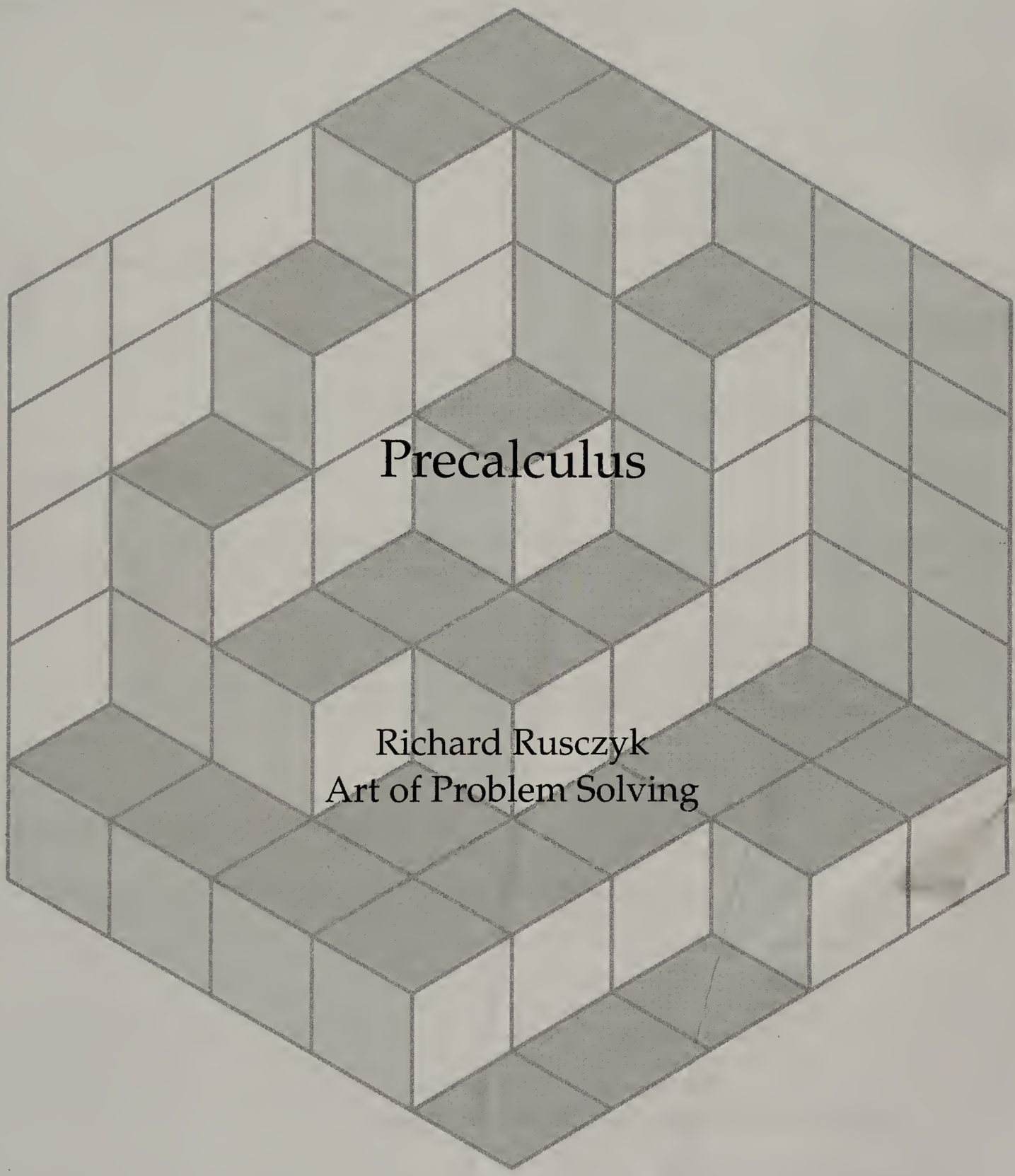
Anno 1778 •

PHILLIPS • ACADEMY

OLIVER•WENDELL•HOLMES

LIBRARY

Per ampliora ad altiora.



Precalculus

Richard Rusczyk
Art of Problem Solving

© 2009, AoPS Incorporated. All Rights Reserved.

Reproduction of any portion of this book without the written permission of AoPS Incorporated is strictly prohibited, except for "fair use" or other noncommercial uses as defined in Sections 107 and 108 of the U.S. Copyright Act.

Published by: AoPS Incorporated
P.O. Box 2185
Alpine, CA 91903-2185
(619) 659-1612
books@artofproblemsolving.com

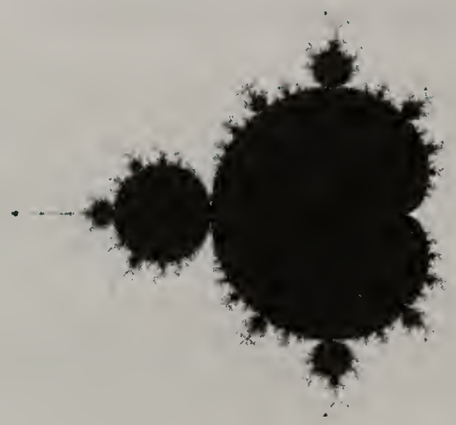
ISBN-13: 978-1-934124-16-1

Visit the Art of Problem Solving website at http://www.artofproblemsolving.com

Cover image designed by Vanessa Rusczyk using KaleidoTile software.

Printed in the United States of America.

First Printing 2009.



How to Use This Book

Organization

The first chapter of this text is a review of functions. Students who have studied our *Introduction to Algebra* and *Intermediate Algebra* textbooks can safely skim this material. Students who are less familiar with functions should study this chapter closely.

The rest of the book is divided into three parts. Chapters 2 through 5 are a thorough exploration of trigonometric functions, trigonometric identities, and their application to Euclidean and analytic geometry. Chapters 6 through 8 cover complex numbers and their relationships to trigonometry and geometry. Chapters 9 through 12 offer an introduction to vectors and matrices in two and three dimensions, and in chapter 13 we apply vectors to Euclidean geometry.

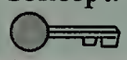
Learn by Solving Problems

We believe that the best way to learn mathematics is by solving problems. Lots and lots of problems. In fact, we believe that the best way to learn mathematics is to try to solve problems that you don't know how to do. When you discover something on your own, you'll understand it much better than if someone had just told you.

Most of the sections of this book begin with several problems. The solutions to these problems will be covered in the text, but try to solve the problems *before* reading the section. If you can't solve some of the problems, that's OK, because they will all be fully solved as you read the section. Even if you can solve all of the problems, it's still important to read the section, both to make sure that your solutions are correct, and also because you may find that the book's solutions are simpler or easier to understand than your own.

Explanation of Icons

Throughout the book, you will see various shaded boxes and icons.

Concept:  This will be a general problem-solving technique or strategy. These are the "keys" to becoming a better problem solver!

Important: This will be something important that you should learn. It might be a formula, a solution technique, or a caution.

WARNING!! Beware if you see this box! This will point out a common mistake or pitfall.

Sidenote: This box will contain material that, although interesting, is not part of the main material of the text. It's OK to skip over these boxes, but if you read them, you might learn something interesting!

Bogus Solution: Just like the impossible cube shown to the left, there's something wrong with any "solution" that appears in this box.

Exercises, Review Problems, and Challenge Problems

Most sections end with several **Exercises**. These will test your understanding of the material that was covered in the section. You should try to solve *all* of the exercises. Exercises marked with a ⋆ are more difficult.

Each chapter concludes with many **Review Problems** and **Challenge Problems**. The Review Problems test your basic understanding of the material covered in the chapter. Your goal should be to solve most or all of the Review Problems for every chapter—if you're unable to do this, it means that you haven't yet mastered the material, and you should probably go back and read the chapter again.

The Challenge Problems are generally more difficult than the other problems in the book, and will really test your mastery of the material. Some of them are very hard—the hardest ones are marked with a ⋆. Don't expect to be able to solve all of the Challenge Problems on your first try—these are difficult problems even for experienced problem solvers. If you are able to solve a large number of Challenge Problems, then congratulations, you are on your way to becoming an expert problem solver!

Hints

Many problems come with one or more hints. You can look up the hints in the Hints section in the back of the book. The hints are numbered in random order, so that when you're looking up a hint to a problem you don't accidentally glance at the hint to the next problem at the same time.

It is very important that you first try to solve each problem without resorting to the hints. Only after you've seriously thought about a problem and are stuck should you seek a hint. Also, for problems which have multiple hints, use the hints one at a time; don't go to the second hint until you've thought about the first one.

Solutions

The solutions to all of the Exercises, Review Problems, and Challenge Problems are in the separate solutions book. If you are using this textbook in a regular school class, then your teacher may decide not to make this solutions book available to you, and instead present the solutions him/herself. However, if you are using this book to learn on your own, then you probably have a copy of the solutions book, in which case there are some very important things to keep in mind:

1. Make a serious attempt to solve each problem before looking at the solution. Don't use the solutions book as a crutch to avoid really thinking about the problem first. You should think *hard* about a problem before deciding to look at the solution. On the other hard, after you've thought hard about a problem, don't feel bad about looking at the solution if you're really stuck.
2. After you solve a problem, it's usually a good idea to read the solution, even if you think you know how to solve the problem. The solutions book might show you a quicker or more concise way to solve the problem, or it might have a completely different solution method than yours.
3. If you have to look at the solution in order to solve a problem, make a note of that problem. Come back to it in a week or two to make sure that you are able to solve it on your own, without resorting to the solution.

Resources

Here are some other recommended resources for you to further pursue your study of mathematics:

- Other Art of Problem Solving textbooks. This series of texts forms a full mathematics curriculum for high-performing students in grades 7-12.
- *The Art of Problem Solving* books, by Sandor Lehoczky and Richard Rusczyk. Whereas the book that you're reading right now will go into great detail of a narrow set of subjects, *the Art of Problem Solving* books cover a wide range of problem solving topics across many different areas of mathematics.
- The `www.artofproblemsolving.com` website. The publishers of this book are also the webmasters of the Art of Problem Solving website, which contains many resources for students:
  - a discussion forum
  - online classes
  - resource lists of books, contests, and other websites
  - a LaTeX tutorial
  - a math and problem solving Wiki
  - and much more!
- You can hone your problem solving skills (and perhaps win prizes!) by participating in various math contests. More details about some of these contests are on page vii, and links to these and many other contests are available on the Art of Problem Solving website.

A Note to Teachers

We believe students learn best when they are challenged with hard problems that at first they may not know how to do. This is the motivating philosophy behind this book.

Rather than first introducing new material and then giving students exercises, we present problems at the start of each section that students should try to solve *before* new material is presented. The goal is to get students

to discover the new material on their own. Often, complicated problems are broken into smaller parts, so that students can discover new techniques one piece at a time. After the problems, new material is formally presented in the text, and full solutions to each problem are explained, along with problem-solving strategies.

We hope that teachers will find that their stronger students will discover most of the material in this book on their own by working through the problems. Other students may learn better from a more traditional approach of first seeing the new material, then working the problems. Teachers have the flexibility to use either approach when teaching from this book.

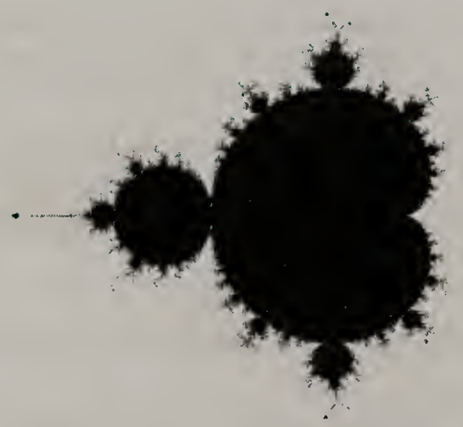
Sections marked with a $\star$ contain supplementary material that may be safely skipped. In general, chapters are not equal in length, so different chapters may take different amounts of classroom time.

Links

The Art of Problem Solving website has a page containing links to websites with content relating to material in this book, as well as an errata list for the book. This page can be found at:

`http://www.artofproblemsolving.com/BookLinks/Precalculus/links.php`

Extra! Occasionally, you'll see a box like this at the bottom of a page. This is an "Extra!" and might be a quote, some biographical or historical background, or perhaps an interesting idea to think about.



Acknowledgements

Contests

We would like to thank the following contests for allowing us to use a wide selection of their problems in this book:

- The **American Mathematics Competitions**, a series of contests for U.S. middle and high school students. The **AMC 8**, **AMC 10**, and **AMC 12** contests are multiple-choice tests that are taken by over 400,000 students every year. Top scorers on the AMC 10 and AMC 12 are invited to take the **American Invitational Mathematics Examination (AIME)**, which is a more difficult, short-answer contest. Approximately 10,000 students every year participate in the AIME. Then, based on the results of the AMC and AIME contests, about 500 students are invited to participate in the **USA Mathematical Olympiad (USAMO)**, a 2-day, 9-hour examination in which each student must show all of his or her work. Results from the USAMO are used to invite a number of students to the Math Olympiad Summer Program, at which the U.S. team for the **International Mathematical Olympiad (IMO)** is chosen. More information about the AMC contests can be found on the AMC website at `www.unl.edu/amc`. (Note: Problems marked **AHSME** are from the American High School Math Exam, the predecessor of the AMC 12.)

- The **Mandelbrot Competition**, which was founded in 1990 by Sandor Lehoczky, Richard Rusczyk, and Sam Vandervelde. The aim of the Mandelbrot Competition is to provide a challenging, engaging mathematical experience that is both competitive and educational. Students compete both as individuals and in teams. The Mandelbrot Competition is offered at the national level for more advanced students and at the regional level for less experienced problem solvers. More information can be found at `www.mandelbrot.org`.

- The **USA Mathematical Talent Search (USAMTS)**, which was founded in 1989 by Professor George Berzsenyi. The USAMTS is a free mathematics competition open to all United States middle and high school students. As opposed to most mathematics competitions, the USAMTS allows students a full month to work out their solutions. Carefully written justifications are required for each problem. More information is available at `www.usamts.org`.

- The **American Regions Math League (ARML)**, which was founded in 1976. The annual ARML competition brings together nearly 2,000 of the nation's finest students. They meet, compete against, and socialize with one another, forming friendships and sharpening their mathematical skills. The contest is written for high school students, although some exceptional junior high students attend each year. The competition consists of several events, which include a team round, a power question (in which a team solves proof-oriented

questions), an individual round, and two relay rounds. More information is available at `www.arml.com`. Some questions from the **New York State Math League (NYSML)** also appear in this text. NYSML is a New York statewide competition that inspired the start of ARML.

- The **Harvard-MIT Mathematics Tournament (HMMT)**, which is an annual math tournament for high school students, held at MIT and at Harvard in alternating years. It is run exclusively by MIT and Harvard students, most of whom themselves participated in math contests in high school. More information is available at `web.mit.edu/hmmt/`.

Some other sources for problems in the book are listed below:

- Mathematical journals *Mathematics & Informatics Quarterly* (M&IQ), *Mathematical Mayhem*, and *Crux Mathematicorum*.
- Various Canadian national competitions, such as the Canadian Invitational Mathematics Contest (CIMC), the Canadian Open Mathematics Challenge (COMC), the Canadian Mathematics Olympiad (CMO), and contests run by the Centre for Education in Mathematics and Computing (CEMC).
- Problems from various countries' national math contests. These problems are cited by country name in the text.

How We Wrote This Book

This book was written using the LaTeX document processing system. We thank the authors of the various LaTeX packages that we used while preparing this book, and also the authors of *The LaTeX Companion* for writing a reference book that is not only thorough but also very readable. The diagrams were prepared using METAPOST and Asymptote.

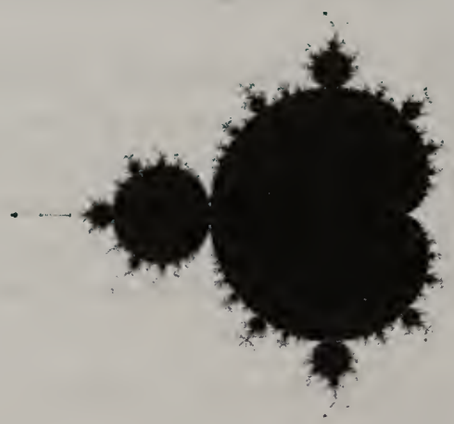
About Us

This book is a collaborative effort of the staff of Art of Problem Solving. Richard Rusczyk was the lead author of the text and primary editor of the solutions manual. Naoki Sato was the lead author of the solutions manual, and selected and organized many of the problems in the text. David Patrick and Jeremy Copeland read several drafts of the text and solutions and made many, many helpful suggestions. Mathew Crawford and Simon Rubinstein-Salzedo collected problems for other Art of Problem Solving books and for Art of Problem Solving classes that are included in this text. The following people also made contributions through proofreading, suggesting material for the textbook, or writing problems for the text: Joshua Brakensiek, James Choi, Billy Dorminy, Gerhardt Hinkle, Patrick Hulin, Vincent Le, Jeff Manning, Aaditya Ramesh, Max Rosett, Sam Ruth, Aaron Schild, Justin Venezuela, and Valentin Vornicu. Vanessa Rusczyk designed the cover and the chapter heading images.

Websites & Errata

We used several websites as source material for the text. Links to these sites and errata for the text and solutions are provided at

`http://www.artofproblemsolving.com/BookLinks/Precalculus/links.php`



Contents

How to Use This Book iii

Acknowledgements vii

1 Functions Review 1

1.1 Function Basics 1
1.2 Graphing Functions 9
1.3 Composition 16
1.4 Inverse Functions 19
1.5 Summary 24

2 Introduction to Trigonometric Functions 29

2.1 The Unit Circle 30
2.2 Radians 40
2.3 Graphs of Trigonometric Functions 45
2.4 Transformations of Trig Functions 55
2.5 Inverse Trig Functions 64
2.6 Summary 76

3 Trigonometric Identities 83

3.1 Introduction to Trig Identities . . . 84
3.2 Sums and Differences of Angles . . . 91
3.3 Double and Half Angles . . . 100
3.4 Sum-to-product and Product-to-sum . . . 106
3.5★ Problem Solving with Identities . . . 112
3.6 Summary . . . 123

4 Applications to Geometry 128

4.1 Right Triangle Trigonometry . . . 128
4.2 Law of Cosines . . . 136
4.3 Law of Sines . . . 141
4.4 More Triangle Relationships . . . 148
4.5 Summary . . . 152

5 Parameterization and Trigonometric Coordinate Systems 158

5.1 Parameterization . . . 158
5.2 Polar Coordinates . . . 167
5.3 Coordinates in Three Dimensions . . . 177
5.4 Summary . . . 187

6 Basics of Complex Numbers 192

6.1 Complex Number Arithmetic . . . 192
6.2 The Complex Plane . . . 198
6.3 Real and Imaginary Parts . . . 202
6.4 Nonreal Roots of Polynomials . . . 208
6.5 Summary . . . 214

7 Trigonometry and Complex Numbers 219

7.1 Polar Form of Complex Numbers . . . 219
7.2 Exponential Form of Complex Numbers . . . 225
7.3 Roots of Unity . . . 233
7.4 Problems Involving Roots of Unity . . . 242
7.5 Summary . . . 250

8 Geometry of Complex Numbers 256

8.1 Transformations of the Complex Plane . . . 256
8.2 Parallel and Perpendicular Lines . . . 265
8.3 Distance . . . 272
8.4 Regular Polygons . . . 278
8.5★ Classic Theorems . . . 286
8.6 Summary . . . 293

9 Vectors in Two Dimensions 298

9.1 What's a Vector? . . . 298
9.2 The Dot Product . . . 305
9.3 Lines and Linear Dependence . . . 313
9.4 Projections . . . 323
9.5 Summary . . . 331

10 Matrices in Two Dimensions 335

10.1 What is a Matrix? . . . 335
10.2 Multiplying Matrices . . . 340
10.3 Matrices as Transformations . . . 348
10.4 The Determinant . . . 355
10.5 Geometric Interpretation of the Determinant . . . 364
10.6 Inverting a Matrix . . . 369
10.7 Summary . . . 375

11 Vectors and Matrices in Three Dimensions, Part 1 380

11.1 Vectors in Three Dimensions . . . 380
11.2 3×3 Matrices . . . 384
11.3 Determinants of 3×3 Matrices . . . 395
11.4 More Than Just 2×2 and 3×3 . . . 409
11.5 Summary . . . 415

12 Vectors and Matrices in Three Dimensions, Part 2 421

12.1 Lines and Planes in Three Dimensions . . . 421
12.2 More Planes in Three Dimensions . . . 428
12.3 The Cross Product . . . 442
12.4 Geometric Interpretation of the Determinant . . . 450

12.5 Inverse of a 3×3 Matrix 454
12.6 Summary 457

13 Vector Geometry **463**
13.1 Introduction 463
13.2 Vectors in the Triangle 469
13.3★ Vectors, Complex Numbers, and Challenging Problems 476
13.4 Summary 485

References **488**

Hints to Selected Problems **489**

Index **501**

For my first two teachers: Mom, who taught me numbers, and Dad, who taught me ways to use them.

Extra! Consider the function

$$f(z) = z^2 + 2i.$$

If we input 0 to this function, we have $f(0) = 0^2 = 2i$. Now, suppose we take this output and put it back into f. We then have

$$f(f(0)) = f(2i) = (2i)^2 + 2i = -4 + 2i.$$

Similarly, we have

$$\begin{aligned} f^3(0) &= f(f(f(0))) \\ &= f(-4 + 2i) \\ &= (-4 + 2i)^2 + 2i \\ &= (-4)^2 + 2(-4)(2i) + (2i)^2 + 2i \\ &= 12 - 14i. \end{aligned}$$

If we keep going, computing $f^4(0)$, $f^5(0)$, and so on, the magnitudes of the results just get larger and larger. However, if we change our starting function just a little bit, we have a very different result when we compute $f^n(0)$ for larger and larger values of n. For example, suppose we start with the function

$$f(z) = z^2 + \frac{i}{2}.$$

Now, if we compute $f(0)$, $f^2(0)$, $f^3(0)$, and so on, we have

$$\begin{aligned} f(0) &= \frac{i}{2}, \\ f^2(0) &= f\left(\frac{i}{2}\right) = -\frac{1}{4} + \frac{i}{2}, \\ f^3(0) &= f\left(-\frac{1}{4} + \frac{i}{2}\right) = -\frac{3}{16} + \frac{i}{4}, \\ &\vdots \\ f^{20}(0) &\approx -0.1324 + 0.3883i, \\ &\vdots \\ f^{40}(0) &\approx -0.1359 + 0.3930i. \end{aligned}$$

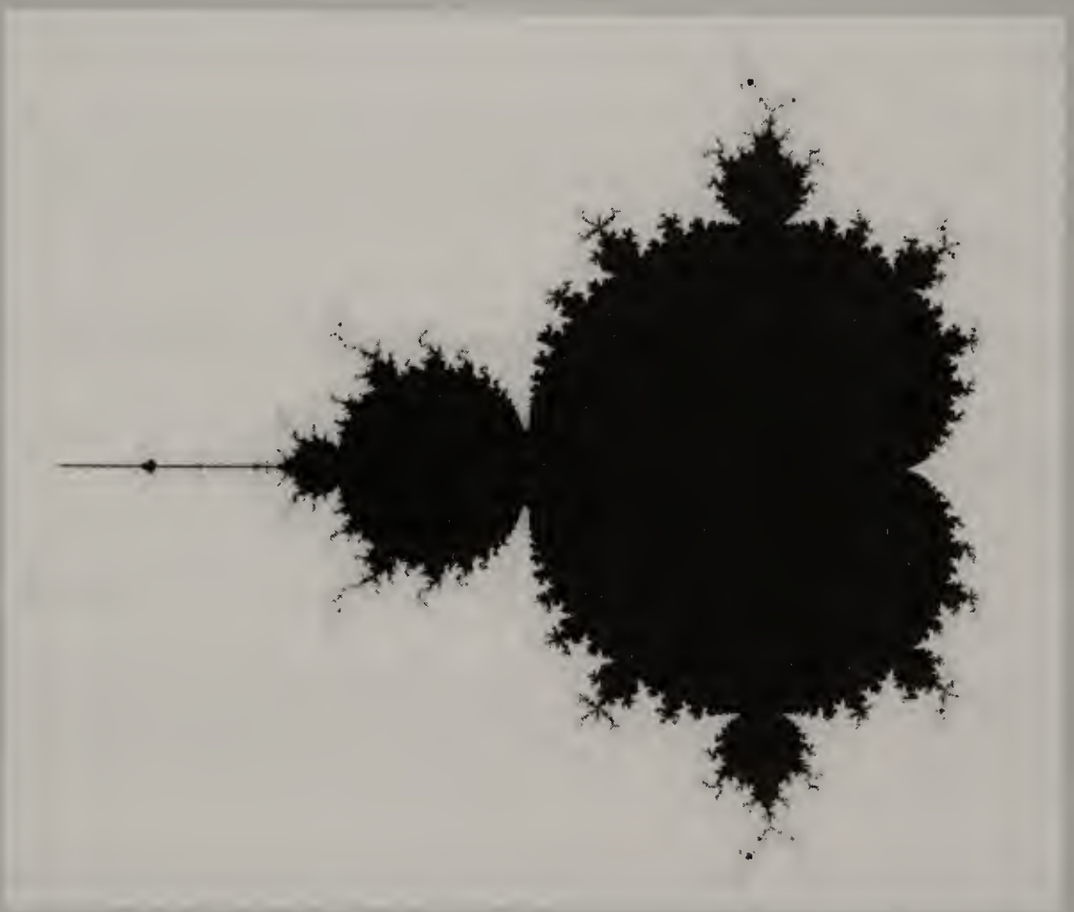
If we keep going and going, we'll find that our results stabilize somewhere around $-0.136 + 0.393i$. The magnitudes of the results don't just keep growing and growing forever! This observation was one of the early breakthroughs in the study of **fractals**.

The only difference between the functions $f(z) = z^2 + 2i$ and $f(z) = z^2 + \frac{i}{2}$ is the constant that is added to z^2 in the function definition. This small change makes a very large difference when we repeatedly compose the function with itself. Of course, $2i$ and $\frac{i}{2}$ aren't the only numbers we can add to z^2 to form our function $f(z)$. With such a big difference between the behavior of $f(z) = z^2 + 2i$ and $f(z) = z^2 + \frac{i}{2}$, we should wonder what happens when we add other complex numbers to z^2.

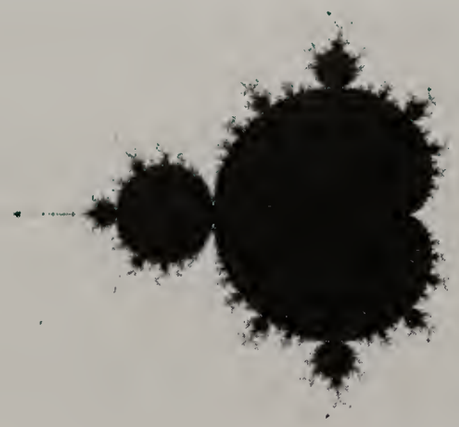
Continued on the following page.

Extra!

We can create a graph based on our results when we try adding different numbers to z^2 in our function definition. We consider all functions of the form $f(z) = z^2 + c$, where c is a constant complex number. For each value of c, we color the point on the complex plane that represents c black or white. If there's no limit to how large the magnitude of $f^n(0)$ gets as n gets larger and larger, then we color the point white. So, our work with $z^2 + 2i$ tells us that the point represented by $c = 2i$ is colored white. If there is a limit to how large the magnitude of $f^n(0)$ gets as n gets larger and larger, then we color the point that represents c black. So, the point that represents $c = \frac{i}{2}$ is colored black.



When we do this coloring for every complex number c on the complex plane, we have a graph that is the geometric depiction of the **Mandelbrot set**, shown above. Unfortunately, low-resolution black-and-white printing cannot do the Mandelbrot set justice. Visit the links page cited on page viii to find links to some fascinating pictures of the Mandelbrot set. If, instead of using white, we color points based on how quickly $f^n(0)$ grows as n grows, the result is quite beautiful. Also, some of these websites contain interactive applets that allow you to zoom in on parts of the Mandelbrot set. With these applets, you'll see some startling images. You'll get a small preview of these images in the chapter headings of this book. The final image in the series of chapter headings is zoomed-in by a factor of over 1,000,000! As you'll see, as we zoom in, we frequently find features of the Mandelbrot set that are *very* similar to the original Mandelbrot set!



The chief function of the body is to carry the brain around. – Thomas Edison

CHAPTER 1

Functions Review

1.1 Function Basics

Suppose we have a machine that accepts any number, multiplies it by two, adds three to this product, and then outputs the result. Mathematically speaking, this machine is a **function** because there is only one possible output from the machine for each input to the machine. We can give this function a label, f, and write the function as

$$f(x) = 2x + 3.$$

This simple equation describes the machine. The "(x)" after f on the left side indicates that we are putting x into the function f. When speaking, $f(x)$ is read "f of x." The x in the equation $f(x) = 2x + 3$ is a **dummy variable**, which means that it is essentially a placeholder. When we put a specific number in our machine, we replace x with that number in the equation $f(x) = 2x + 3$ to determine what the machine outputs. For example,

$$f(5) = 2 \cdot 5 + 3 = 13,$$

so the machine outputs 13 when we put 5 into it.

Functions are really that simple. We define the function, and then whenever we input a number to the function, we follow the definition to get an output. Usually, we use an equation, such as $f(x) = 2x + 3$, to define a function. For the obvious reason, f is the most commonly used label for functions.

In this chapter, we will discuss only functions that take real number inputs and give real number outputs. The **domain** of a function consists of all the values we are able to input to the function and get an output, and the **range** of the function consists of all the values that can come out of the function.

Extra! *It isn't that they can't see the solution. It is that they can't see the problem.*
– G. K. Chesterton

For example, consider the function

$$f(x) = \frac{1}{x-3}.$$

The value $x = 3$ is not part of the domain of this function, because $\frac{1}{x-3}$ is not defined when $x = 3$. We can safely put any other value of x into this function, so the domain of f is "all real numbers except 3." Similarly, there is no value of x for which it is possible to make the function output 0. However, we can make the function output any other real number, so the range of f is "all real numbers except 0."

We call a function that can only output real numbers a **real-valued function**, or, even more simply, a **real function**. Unless a problem states otherwise, you can assume that all functions in this chapter are real functions. Moreover, you can assume that they are only defined for real number inputs. Later in this book we will explore a variety of functions that can take nonreal inputs and give nonreal outputs.

We will sometimes use **interval notation** to express the domain and range of a function. For example, we can denote "all real numbers greater than 3 and less than 5" by the interval $(3, 5)$. We use "(" and ")" to indicate that the values 3 and 5 are not included. Notice that we don't write $(5, 3)$; the lesser end of the interval always comes first. We write $x \in (3, 5)$ to indicate that x is a number in this interval, The "$x \in$" means "x is in," and $(3, 5)$ indicates the real numbers greater than 3 and less than 5, as we just described.

To include a boundary value in the interval, we use "[" for the lower bound and "]" for the upper bound. For example, the statement $x \in (-3, 5]$ means $-3 < x \leq 5$ and $y \in [-12.2, 0]$ means $-12.2 \leq y \leq 0$.

Finally, to indicate an interval that has no upper bound or no lower bound (or neither), we use the ∞ symbol. For example, we write $t > -3$ as $t \in (-3, +\infty)$. The "$(-3$" part indicates that no numbers equal to -3 or lower than -3 are in the interval. The "$+\infty)$" part indicates that the interval continues forever in the positive direction; that is, there is no upper bound. So, the interval $(-3, +\infty)$ is all real numbers greater than -3. Similarly, the statement $w \in (-\infty, -2]$ is the same as $w \leq -2$. Notice that we always use "(" with $-\infty$ and ")" with $+\infty$, instead of "[" and "]". In both cases, there is not a boundary value to include in the interval, so we must use "(" or ")" instead of "[" or "]".

To write all real numbers in interval notation, we can write $(-\infty, +\infty)$. We often use the symbol $\mathbb{R}$ to denote "all real numbers." Therefore, we typically write the interval $(-\infty, \infty)$ as "$\mathbb{R}$", so $x \in \mathbb{R}$ means "x is a real number."

Problems

Problem 1.1: Let $f(x) = 2x^2 - 3$.

(a) Find $f(2)$.

(b) Find all values of t such that $f(t) = 47$.

(c) Find $f(3x + 1)$.

Problem 1.2: A function can be defined to accept multiple inputs. For example, let

$$g(a, b, c) = 3a - 2b + 7c^2.$$

(a) Evaluate $g(3, -5, -1)$.

(b) Find b if $g(b, 2, -1) = 21$.

Problem 1.3: Suppose $t(x) = ax^4 + bx^2 + x + 5$, where a and b are constants. In this problem, we find $t(4)$ if $t(-4) = 3$.

(a) Use $t(-4) = 3$ to write an equation in terms of a and b.

(b) Express $t(4)$ in terms of a and b.

(c) Find $t(4)$.

Problem 1.4: Find the domain and range of each of the following functions:

(a) $f(x) = 2x - 3$

(b) $f(x) = \sqrt{-2x + 7}$

(c) $g(t) = \dfrac{2t}{t-1}$

(d) $h(x) = 9x^2 + 4$

Problem 1.5: Suppose $p(x) = 2\sqrt{x} + 3$, but that $p(x)$ is only defined for $4 \le x \le 9$. In other words, the domain of p is all real numbers from 4 to 9.

(a) Is there a value of x for which $p(x) = 5$?

(b) What is the range of p?

Problem 1.6: Find the domain of each of the following functions:

(a) $f(x) = \dfrac{\sqrt{2x-5}}{x-3}$

(b) $g(t) = \dfrac{2t-4}{\frac{1}{t} - \frac{1}{3t-4}}$

Problem 1.7: Notice that $\dfrac{x^2-x}{x-1} = \dfrac{(x-1)(x)}{x-1} = x$. Are $f(x) = \dfrac{x^2-x}{x-1}$ and $g(x) = x$ the same function?

Problem 1.1: Let $f(x) = 2x^2 - 3$.

(a) Find $f(2)$.

(b) Find all values of t such that $f(t) = 47$.

(c) Find $f(3x + 1)$.

Solution for Problem 1.1:

(a) We simply replace x with 2 in the function definition, which gives us $f(2) = 2(2)^2 - 3 = 8 - 3 = 5$.

(b) If $f(t) = 47$, we must have $2t^2 - 3 = 47$. Adding 3 to both sides gives $2t^2 = 50$, and dividing by 2 gives $t^2 = 25$. The two values of t that satisfy this equation are $t = 5$ and $t = -5$. Checking our answer, we find $f(5) = 2(5)^2 - 3 = 47$ and $f(-5) = 2(-5)^2 - 3 = 47$.

Concept: We can often check our answers in algebra problems by plugging the solutions we find back into the original question.

(c) We can input an entire expression into a function by replacing the dummy variable in the function definition with the expression. Moreover, we can replace the x in the function definition $f(x) = 2x^2 - 3$ with an expression that contains x. For example, we have

$$f(3x + 1) = 2(3x + 1)^2 - 3 = 2(9x^2 + 6x + 1) - 3 = 18x^2 + 12x - 1.$$

□

Problem 1.2: A function can be defined to accept multiple inputs. For example, let

$$g(a, b, c) = 3a - 2b + 7c^2.$$

(a) Evaluate $g(3, -5, -1)$.

(b) Find b if $g(b, 2, -1) = 21$.

Solution for Problem 1.2:

(a) We replace the dummy variables a, b, and c in the function definition with the values 3, -5, and -1, in that order. This gives us $g(3,-5,-1) = 3(3) - 2(-5) + 7(-1)^2 = 26$.

(b) What's wrong with this solution:

Bogus Solution: We have

$$g(b,2,-1) = 3(2) - 2b + 7(-1)^2 = -2b + 13,$$

so we seek the value of b such that $-2b + 13 = 21$, which is $b = -4$.

We see our mistake immediately if we check our answer. We have

$$g(-4,2,-1) = 3(-4) - 2(2) + 7(-1)^2 = -9,$$

but we want $g(b,2,-1) = 21$. Looking at our Bogus Solution above, we see that in finding an expression for $g(b,2,-1)$, we set $a = 2$ and $c = -1$, and left b as b. But this isn't what we must do to evaluate $g(b,2,-1)$.

WARNING!! When evaluating a function that accepts multiple inputs, make sure you assign the values to the dummy variables in the correct order.

To evaluate $g(b,2,-1)$, we replace a in the function definition with b, we replace b with 2, and replace c with -1. This gives us

$$g(b,2,-1) = 3b - 2(2) + 7(-1)^2 = 3b + 3.$$

So, we seek the value of b such that $3b + 3 = 21$, which gives us $b = 6$. Checking our answer, we find that $g(6,2,-1) = 21$, as desired.

□

Problem 1.3: Suppose $t(x) = ax^4 + bx^2 + x + 5$, where a and b are constants. Find $t(4)$ if $t(-4) = 3$.

Solution for Problem 1.3: We have

$$t(4) = a(4)^4 + b(4)^2 + 4 + 5 = 256a + 16b + 4 + 5 = 256a + 16b + 9.$$

If we can find a and b, then we can evaluate $t(4)$. We turn to the only other piece of information we have, which is $t(-4) = 3$. We have

$$t(-4) = a(-4)^4 + b(-4)^2 + (-4) + 5 = 256a + 16b - 4 + 5 = 256a + 16b + 1,$$

so $t(-4) = 3$ gives us $256a + 16b + 1 = 3$, which means $256a + 16b = 2$. Unfortunately, this doesn't tell us a or b. However, looking back at our expression for $t(4)$, we see that we don't need a and b. We need $256a + 16b$, which we know equals 2. So, we have $t(4) = 256a + 16b + 9 = 2 + 9 = 11$. □

Concept: Keep your eye on the ball. Sometimes we don't need to evaluate every variable in a problem in order to solve the problem.

Now that we have a little practice evaluating functions, let's try finding the domain and range of specific functions.

Problem 1.4: Find the domain and range of each of the following functions:

(a) $f(x) = 2x - 3$

(b) $f(x) = \sqrt{-2x + 7}$

(c) $g(t) = \dfrac{2t}{t-1}$

(d) $h(x) = 9x^2 + 4$

Solution for Problem 1.4:

(a) We can input any real number to f, so the domain of f is all real numbers. It appears that any real number can be output from f. To show that the range of f is all real numbers, we show that for any real number y, we can find an x such that $f(x) = y$. Since $f(x) = 2x - 3$, we seek the value of x such that $2x - 3 = y$. Solving for x gives $x = (y + 3)/2$. So, for any real number y, if we let $x = (y + 3)/2$, then we have $f(x) = y$. This tells us that every real number is in the range of f. Therefore, both the domain and range are all real numbers. Since we can use the symbol $\mathbb{R}$ to indicate the real numbers, we can write, "The domain and range of f are both $\mathbb{R}$."

(b) The square root of a negative number is not a real number, so we must have $-2x + 7 \geq 0$. This gives us $x \leq 7/2$, so the domain is all real numbers less than or equal to $7/2$. In interval notation, the domain is $(-\infty, 7/2]$.

Intuitively, we might guess that because $-2x + 7$ can equal any nonnegative number, the range of f is all nonnegative real numbers. We can explicitly show that this is the range of f as we did in part (a). We let

$$y = f(x) = \sqrt{-2x + 7}.$$

We then solve this equation for x in terms of y. Squaring both sides of $y = \sqrt{-2x + 7}$ gives $y^2 = -2x + 7$. We solve for x by subtracting 7 from both sides, then dividing by -2, to find

$$x = \frac{7 - y^2}{2}.$$

So, for any nonnegative value of y, if we let $x = (7 - y^2)/2$, then we have

$$f(x) = f\left(\frac{7 - y^2}{2}\right) = \sqrt{-2\left(\frac{7 - y^2}{2}\right) + 7} = \sqrt{-7 + y^2 + 7} = \sqrt{y^2} = y.$$

This final step, $\sqrt{y^2} = y$, is only valid *because y is nonnegative.* We therefore see that all nonnegative real numbers are in the range of f, because for every nonnegative value of y, we can find an x such that $f(x) = y$. Since $\sqrt{-2x + 7}$ cannot be negative, no negative numbers are in the range of f. So, the range of f is all nonnegative real numbers, which we can write as $[0, +\infty)$.

(c) We cannot have $t = 1$ in $g(t) = 2t/(t - 1)$, since this would make the denominator equal to 0. There are no other restrictions on the input to g, so the domain is all real numbers except 1.

Finding the range of g is a little trickier. We let $y = g(t)$, so we have

$$y = \frac{2t}{t - 1}.$$

Then we solve for t in terms of y. Multiplying both sides by $t - 1$ gives $yt - y = 2t$. Solving this equation for t in terms of y gives

$$t = \frac{y}{y - 2}.$$

(Note that t cannot be 1 in this equation, since $t = 1$ gives us $1 = \frac{y}{y-2}$, from which we have $y - 2 = y$, which has no solution for y.) For any desired output y except $y = 2$, we can use $t = \frac{y}{y-2}$ to find the input, t, to $g(t)$ that will produce the desired output. Therefore, the range is all real numbers except 2. We can denote "all real numbers except 2" with interval notation as $(-\infty, 2) \cup (2, +\infty)$.

The "$\cup$" in $(-\infty, 2) \cup (2, +\infty)$ means "or," so, "$y \in (-\infty, 2) \cup (2, +\infty)$" means

$$y \text{ is in the interval } (-\infty, 2) \text{ or the interval } (2, +\infty).$$

Sidenote:

The **union** of two intervals consists of all numbers that are in either one or both of the intervals. As we just saw, we use the symbol $\cup$ to denote a union of two intervals, so $(-\infty, 2) \cup (2, +\infty)$ means "all numbers in the interval $(-\infty, 2)$ or the interval $(2, +\infty)$."

The **intersection** of two intervals consists of all numbers that are in both of the intervals. We use the symbol $\cap$ to refer to the intersection of two intervals. For example, the expression $[3, 7] \cap [5, 11]$ refers to the numbers that are in both the interval $[3, 7]$ *and* the interval $[5, 11]$. These numbers form the interval $[5, 7]$, so we can write

$$[3, 7] \cap [5, 11] = [5, 7].$$

On the other hand, the numbers that are either in $[3, 7]$ or in $[5, 11]$ (or in both) form the interval $[3, 11]$. Therefore, we have

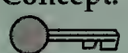
$$[3, 7] \cup [5, 11] = [3, 11].$$

We also have a special notation for "all real numbers except a few specific values." We can write "all real numbers except 2" as $\mathbb{R}\backslash\{2\}$. If we wish to exclude several specific values, we just list them inside the curly braces. So, we write "all real numbers except 5, 6, and 7" as $\mathbb{R}\backslash\{5, 6, 7\}$.

(d) We can input any real number to the function $h(x) = 9x^2 + 4$, so the domain of h is $\mathbb{R}$. Because the square of a real number is always nonnegative, the expression $9x^2$ is always nonnegative. Therefore, the expression $9x^2 + 4$ must be greater than or equal to 4. Since $9x^2 + 4$ can equal any number that is greater than or equal to 4, the range of h is $[4, +\infty)$.

□

Note that in parts (b) and (c), we find the domain by considering operations that we are not able to perform. Specifically, we cannot take the square root of a negative number in a real function, and we cannot divide by zero. This is typically how we find the domain when it is not obvious that the domain is all real numbers, as it is in parts (a) and (d). Finding the range is often a matter of considering the possible outputs of special expressions like square roots, perfect squares, or absolute values. But, as shown in part (c), the restrictions on the range can also be more subtle. There, we found a strategy that is often useful for determining the range of complicated functions:

Concept: 

We can often find the range of a function by setting the function definition equal to a new variable, and then solving for the dummy variable in terms of the new variable.

For some functions, we need the advanced tools of calculus to determine the range. For still others, the best we can do is estimate the range numerically, usually using a computer. When we ask for the range of a function in this text, we'll only consider functions whose ranges we can find without these techniques.

Sometimes functions are defined with explicit constraints on the domain of the function. The range of such a function consists only of those outputs that can be obtained from the permitted inputs. Here's an example:

Problem 1.5: Suppose $p(x) = 2\sqrt{x} + 3$, but that $p(x)$ is only defined for $4 \leq x \leq 9$. (In other words, the domain of p is all real numbers from 4 to 9.) Find the range of p.

Solution for Problem 1.5: What's wrong with this solution:

Bogus Solution:

The output of $\sqrt{x}$ can be any nonnegative real number, so the result of $2\sqrt{x} + 3$ can be any nonnegative real number greater than or equal to 3.

Our Bogus Solution completely ignores the restriction on the domain of p.

WARNING!! We must take into account any restrictions on the domain of a function when finding the range of the function.

Since p is only defined for inputs x such that $4 \le x \le 9$, the range consists of all values of p that can result from inputting a value of x such that $4 \le x \le 9$. Because $4 \le x \le 9$, we have $2 \le \sqrt{x} \le 3$. Multiplying all parts of this inequality chain by 2 gives $4 \le 2\sqrt{x} \le 6$ and adding 3 to all parts of this chain gives $7 \le 2\sqrt{x} + 3 \le 9$. Therefore, for all x such that $4 \le x \le 9$, we have $7 \le p(x) \le 9$.

We're not finished yet! We've only shown that all values in the range of p are between 7 and 9, inclusive (which means 7 and 9 are included). We haven't shown that all real numbers from 7 to 9 are in the range. This is somewhat "obvious," since $p(x)$ goes "smoothly" from 7 to 9 as x goes from 4 to 9. However, to explicitly prove that all possible values from 7 to 9 are in the range, we let $y = p(x) = 2\sqrt{x} + 3$ and solve for x in terms of y. This gives us

$$x = \left(\frac{y-3}{2}\right)^2.$$

For any value of y from 7 to 9, we can use this equation to find the value of x for which $p(x)$ equals this y. (Make sure you see why all the resulting values of x are in the domain of p.) □

Finding the domain and range of more complicated functions can be pretty tricky. We'll try finding the domain of a couple of such functions.

Problem 1.6: Find the domain of each of the following functions:

(a) $f(x) = \dfrac{\sqrt{2x-5}}{x-3}$

(b) $g(t) = \dfrac{2t-4}{\frac{1}{t} - \frac{1}{3t-4}}$

Solution for Problem 1.6:

(a) The input to a square root must be nonnegative, so we must have $2x-5 \ge 0$. This gives us $x \ge 5/2$. However, we cannot have $x = 3$, since this will make the denominator equal to 0.

WARNING!! Sometimes there is more than one constraint on the domain of a function.

Combining these two constraints tells us that the domain is all real numbers greater than or equal to 5/2 except 3. We can write this in interval notation as $[5/2, 3) \cup (3, +\infty)$.

(b) First, we notice that we cannot have $t = 0$ or $3t - 4 = 0$, because we cannot divide by 0. This tells us that we must exclude $t = 0$ and $t = 4/3$ from the domain. However, we're not finished, because we cannot have the entire denominator of $g(t)$ equal to 0. Therefore, we must exclude from our domain those values of t for which

$$\frac{1}{t} - \frac{1}{3t-4} = 0.$$

Multiplying both sides by $t(3t-4)$ to get rid of the fractions gives us $3t - 4 - t = 0$, from which we find $t = 2$. So, the domain of g is all real numbers except 0, 4/3, and 2. We could write this as intervals with the unwieldy $(-\infty, 0) \cup (0, \frac{4}{3}) \cup (\frac{4}{3}, 2) \cup (2, +\infty)$. We can also write it as $\mathbb{R}\backslash\{0, \frac{4}{3}, 2\}$, which reads "all real numbers except 0, $\frac{4}{3}$, and 2."

□

Problem 1.7: Notice that $\dfrac{x^2-x}{x-1} = \dfrac{(x-1)(x)}{x-1} = x$. Are $f(x) = \dfrac{x^2-x}{x-1}$ and $g(x) = x$ the same function?

Solution for Problem 1.7: No! The function f is not defined for $x = 1$, even though it appears that we can simplify our expression for $f(x)$ to just x. The simplification

$$\frac{x^2-x}{x-1} = \frac{x(x-1)}{x-1} = x$$

is only valid if $x \neq 1$. If $x = 1$, then $\frac{x^2-x}{x-1}$ is $\frac{0}{0}$, which is not defined. In other words, 1 is not in the domain of f, but it is in the domain of g. So, f and g are not the same function. □

We finish this section with the relatively obvious observation that the sum, product, and quotient of two functions are all themselves functions. For example, if f and g are functions, we can define a new function h as $h = f + g$, where this means that $h(x) = f(x) + g(x)$. Note that $h(x)$ is only defined if both $f(x)$ and $g(x)$ are defined. Therefore, a number must be in the domains of both f and g in order to be in the domain of $f + g$.

In much the same way, we can define a new function as a linear combination of any two functions f and g. Specifically, if a and b are constants, then we can define a new function t as $t = a \cdot f + b \cdot g$, where this means that $t(x) = a \cdot f(x) + b \cdot g(x)$.

We can also define the function $p = f \cdot g$ such that $p(x) = f(x) \cdot g(x)$ and the function $q = f/g$ such that $q(x) = f(x)/g(x)$. Note that if $q = f/g$, then we must exclude values of x for which $g(x) = 0$ from the domain of q.

Exercises

1.1.1 Let $f(x) = x^2 - x - 6$.

(a) Compute $f(2)$.

(b) Find all x such that $f(x) = 0$.

(c) Find $f(x-1)$.

(d) Is there a real value of x such that $f(x) = 6$?

(e) Find the range of f. **Hints:** 1

1.1.2 Find the domain and range of each of the following real-valued functions:

(a) $f(x) = |1 - x|$ (b) $f(t) = \sqrt{2-t}$ (c) $h(x) = 1 - x^2$ (d) $g(u) = \frac{1}{1+\frac{1}{u}}$

1.1.3 Let $f(x) = \sqrt{4-x}$ and $g(x) = \sqrt{2x-6}$. Find the domain of $f \cdot g$ and the domain of f/g. Explain why these domains are different.

1.1.4 Do the functions $f(x) = \sqrt{\dfrac{2x-5}{x-8}}$ and $g(x) = \dfrac{\sqrt{2x-5}}{\sqrt{x-8}}$ have the same domain? Why or why not?

1.1.5 Let $f(x) = \dfrac{x+1}{x-1}$. Suppose $|x| \neq 1$. Simplify $f(x) \cdot f(-x)$.

1.1.6 A function $g(x)$ is defined for $-3 \leq x \leq 4$ such that $g(x) = (2+x)^2$. Find the range of g.

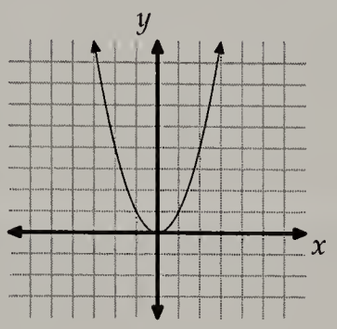
1.1.7 Let $T(a,b,c) = 3a^b - c$.

(a) Find $T(2,3,-5)$.

(b) Find all values of x such that $T(x,2,6) = 21$.

1.2 Graphing Functions

When we **graph an equation** in which the only variables are x and/or y, we plot all points (x, y) on the Cartesian plane that satisfy the equation. When we **graph a function** f, we graph the equation $y = f(x)$. For example, the graph of the function $f(x) = x^2$ is shown at right.



The **x-intercepts** of a graph are the points where the graph crosses the x-axis, while the **y-intercepts** are the points where the graph crosses the y-axis. For example, the point $(0, 0)$ is both an x-intercept and a y-intercept of the graph of $f(x) = x^2$. (Note: Some sources use "x-intercept" to refer to the x-coordinate of a point where a graph crosses the x-axis, and likewise for "y-intercept." So, these sources would describe both the x-intercept and the y-intercept of the graph of $f(x) = x^2$ as 0 instead of $(0, 0)$.)

In this section, we review the basics of graphing functions, getting information from the graph of a function, and transforming the graph of a function.

Problems

Problem 1.8: For each of the functions below, graph the function and find the x-intercepts and the y-intercepts of the graph.

(a) $f(x) = (x - 6)^2 - 4$

(b) $f(x) = 4$

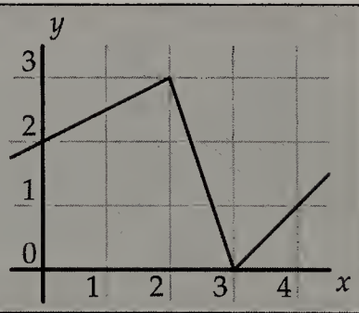
(c) $f(x) = \sqrt{x + 5} - 3$

(d) $f(x) = |x + 2| + 1$

Problem 1.9: The graph of $y = f(x)$ is pictured at right.

(a) According to the graph, what is $f(2)$?

(b) According to the graph, for what values of a is $f(a) = 2$?



Problem 1.10:

(a) Is it possible for a vertical line to intersect the graph of a function at two different points? Why or why not?

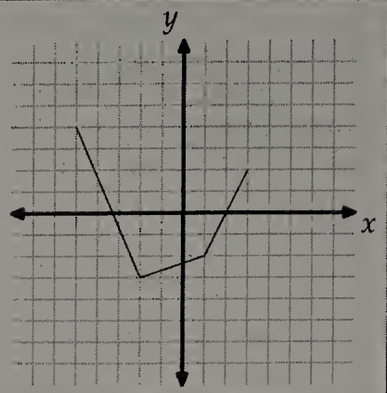
(b) If there is no vertical line that passes through more than one point of a curve on the Cartesian plane, then is the curve the graph of a function?

Problem 1.11: At right is the graph of $y = f(x)$.

(a) Explain why the graph of $y = f(x) + 2$ is a 2-unit upward shift of the graph of $y = f(x)$.

(b) Explain why the graph of $y = f(x + 2)$ is a 2-unit leftward shift of the graph of $y = f(x)$.

(c) Find the domain and range of h if $h(x) = f(x + 3) - 4$.



Problem 1.12: Let $f(x)$ be the function graphed in the previous problem.

(a) Find the graph of $y = -f(x)$.

(b) Find the graph of $y = f(-x)$.

(c) Find the graph of $y = 2f(x)$.

(d) Find the graph of $y = f(2x)$.

Problem 1.8: For each of the functions below, graph the function and find the x-intercepts and the y-intercepts of the graph.

(a) $f(x) = (x - 6)^2 - 4$

(b) $f(x) = 4$

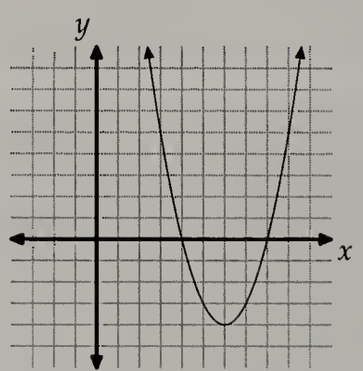
(c) $f(x) = \sqrt{x + 5} - 3$

(d) $f(x) = |x + 2| + 1$

Solution for Problem 1.8:

(a) The graph of $y = (x - 6)^2 - 4$ is an upward-opening parabola with vertex $(6, -4)$. We find several points on the graph by finding y for various values of x. A table with these points is at left below, and the graph is at right below.

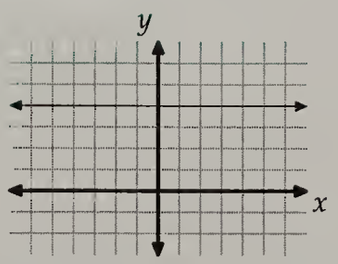
x	$y = f(x)$
3	5
4	0
5	−3
6	−4
7	−3
8	0
9	5



The y-intercept of the graph is the point where the graph intersects the y-axis. All points on the y-axis have an x-coordinate of 0, so evaluating $f(0)$ gives us the y-coordinate of the y-intercept of the graph of $y = f(x)$. Since $f(0) = (0 - 6)^2 - 4 = 32$, the y-intercept of the graph is $(0, 32)$.

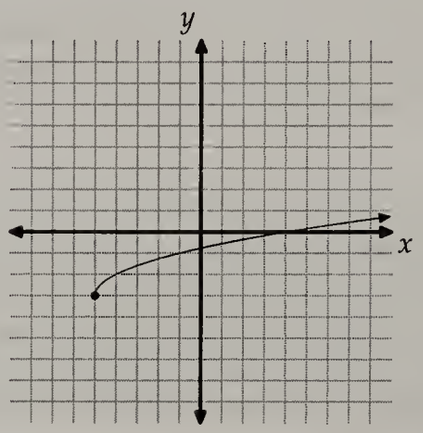
The x-intercepts of the graph are those points where the graph intersects the x-axis. The y-coordinate of such a point is 0, so the x-coordinates of the x-intercepts are the solutions to the equation $f(x) = 0$. Therefore, we seek the values of x such that $(x - 6)^2 - 4 = 0$. Expanding and rearranging gives us $x^2 - 12x + 32 = 0$. Factoring gives $(x - 4)(x - 8) = 0$, so our solutions are $x = 4$ and $x = 8$. Therefore, the x-intercepts of our graph are $(4, 0)$ and $(8, 0)$. (Notice that we could also read these off our graph.)

(b) Since $f(x) = 4$, the graph of $y = f(x)$ is the graph of $y = 4$, which is a horizontal line as shown at right. This line meets the y-axis at $(0, 4)$, so this is the y-intercept. The graph never hits the x-axis, so it does not have an x-intercept.



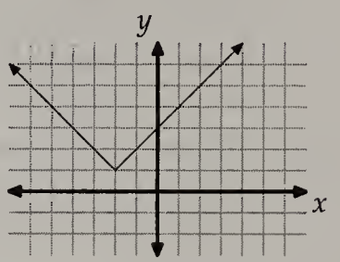
(c) As with the first part, we choose values of x, and then find y such that $y = f(x)$. Notice that we start with $x = -5$ and then substitute greater values of x, because the domain of f is $[-5, +\infty)$.

x	$y = f(x)$
-5	-3
-4	-2
-1	-1
4	0
11	1



When $x = 0$, we have $y = f(0) = -3 + \sqrt{5}$, so the y-intercept is $(0, -3 + \sqrt{5})$. When $y = 0$, we have $0 = \sqrt{x+5} - 3$. Adding 3 to both sides gives $3 = \sqrt{x+5}$, and squaring both sides of this gives $9 = x + 5$. Solving this gives $x = 4$, so $(4, 0)$ is the x-intercept of the graph.

(d) If $x \geq -2$, we have $|x + 2| = x + 2$, so if $x \geq -2$, we have $f(x) = x + 2 + 1 = x + 3$. So, when graphing $y = f(x)$, we graph $y = x + 3$ for $x \geq -2$. Similarly, if $x < -2$, we have $|x + 2| = -x - 2$, so graphing $y = f(x)$ for $x < -2$ is the same as graphing $y = -x - 2 + 1 = -x - 1$. The result is shown at right.



When $x = 0$, we have $y = f(0) = 3$, so $(0, 3)$ is the y-intercept of the graph. When $y = 0$, we have $0 = |x + 2| + 1$. Subtracting 1 gives $|x + 2| = -1$, which has no solutions because absolute value is always nonnegative. So, the graph has no x-intercept. (We can also see this from our graph.)

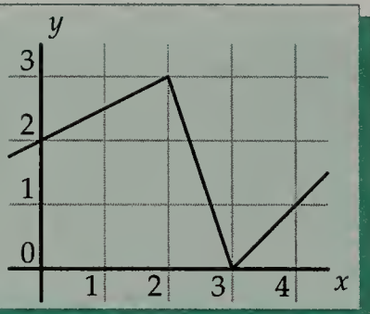
□

Now that we've used the definition of a function to get information about the graph of the function, let's try going the other direction—using a graph to get information about the function.

Problem 1.9: The graph of $y = f(x)$ is pictured at right.

(a) According to the graph, what is $f(2)$?

(b) According to the graph, for what values of a is $f(a) = 2$?



Solution for Problem 1.9:

(a) The graph of $y = f(x)$ passes through $(2, 3)$, so we have $f(2) = 3$.

Important: If the point (a, b) is on the graph of $y = f(x)$, then $f(a) = b$. Conversely, if $f(a) = b$, then (a, b) is on the graph of $y = f(x)$.

(b) To determine the values of a such that $f(a) = 2$, we locate the points on the graph of $y = f(x)$ such that $y = 2$. The x-coordinates of these points then are our desired values of a. Since $(0, 2)$ is on the graph of $y = f(x)$, we have $f(0) = 2$.

There is another point on the graph with y-coordinate equal to 2, but it's not immediately obvious what the x-coordinate of this point is. However, this point is on the line segment connecting $(2, 3)$ to $(3, 0)$, so we can use the equation of the line through these two points to find the desired x-coordinate. The slope of this line is $(3 - 0)/(2 - 3) = -3$, so a point-slope equation of the line is $y - 0 = -3(x - 3)$, or $y = -3x + 9$. When $y = 2$ in this equation, we have $x = 7/3$, so the line passes through $(7/3, 2)$. Therefore, we have $f(7/3) = 2$, so the two values of a for which $f(a) = 2$ are $a = 0$ and $a = 7/3$.

□

Problem 1.10:

(a) Is it possible for a vertical line to intersect the graph of a function at two different points? Why or why not?

(b) If there is no vertical line that passes through more than one point of a curve on the Cartesian plane, then is the curve the graph of a function?

Solution for Problem 1.10:

(a) For each value of x in the domain of f, there is exactly one value of $f(x)$. So, for each such value of x, there is only one point (x, y) on the graph of $y = f(x)$. Therefore, it is impossible for the graph of a function to pass through two points with the same x-coordinate. Because the equation of any vertical line has the form $x = a$ for some constant a, it is impossible for a vertical line to pass through more than one point on the graph of a function.

(b) If there is not a vertical line that passes through more than one point of a graph, then for every value of x, there is at most one value of y such that (x, y) is on the graph. In other words, for every x for which there is a point (x, y) on the graph, there is exactly one corresponding value of y, so the graph does indeed represent a function.

□

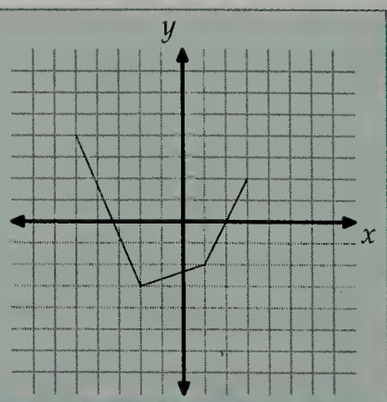
Important: A curve on the Cartesian plane is the graph of some function if and only if every vertical line passes through no more than one point on the curve. We call this the **vertical line test**.

Problem 1.11: At right is the graph of $y = f(x)$.

(a) Explain why the graph of $y = f(x) + 2$ is a 2-unit upward shift of the graph of $y = f(x)$.

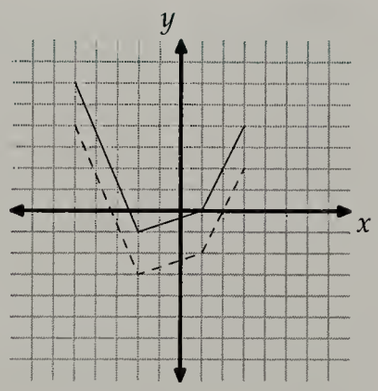
(b) Explain why the graph of $y = f(x + 2)$ is a 2-unit leftward shift of the graph of $y = f(x)$.

(c) Find the domain and range of h if $h(x) = f(x + 3) - 4$.



Solution for Problem 1.11:

(a) Suppose $x = a$, where a is in the domain of f. Then, the point $(a, f(a))$ is on the graph of $y = f(x)$ and the point $(a, f(a)+2)$ is on the graph of $y = f(x)+2$. Therefore, every point of the graph of $y = f(x) + 2$ is 2 units above the corresponding point on the graph of $y = f(x)$. So, the graph of $y = f(x) + 2$ is a 2-unit upward shift of the graph of $y = f(x)$. The graph of $y = f(x) + 2$ is solid in the graph at right, and the graph of $y = f(x)$ is dashed.



Similarly, if k is positive, then the graph of

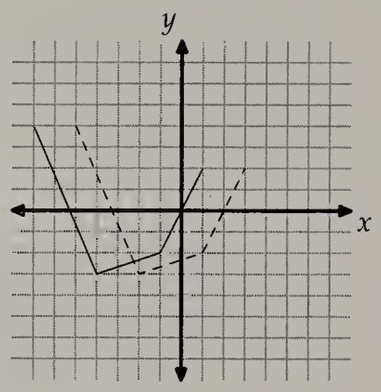
$$y = f(x) + k$$

is a k-unit upward shift of the graph of $y = f(x)$, and the graph of

$$y = f(x) - k$$

is a k-unit downward shift of the graph of $y = f(x)$.

(b) Suppose $x = a$, where a is in the domain of f. Then, the point $(a, f(a))$ is on the graph of $y = f(x)$ and the point $(a, f(a+2))$ is on the graph of $y = f(x+2)$. Unfortunately, it isn't so clear how $f(a)$ and $f(a+2)$ are related. To get a point on the graph of $y = f(x + 2)$ with y-coordinate equal to $f(a)$, we have to let $x = a - 2$. When $x = a - 2$ in $y = f(x + 2)$, we have $y = f(a)$, so $(a - 2, f(a))$ is on the graph of $y = f(x + 2)$.



We've shown that for any a in the domain of f, the point $(a, f(a))$ is on the graph of $y = f(x)$ and the point $(a - 2, f(a))$ is on the graph of $y = f(x + 2)$. Therefore, every point on the graph of $y = f(x+2)$ is 2 units to the left of the corresponding point on the graph of $y = f(x)$. So, the graph of $y = f(x + 2)$ is a 2-unit leftward shift of the graph of $y = f(x)$. The graph of $y = f(x + 2)$ is solid in the graph at right above, and the graph of $y = f(x)$ is dashed.

Similarly, if $k > 0$, then the graph of

$$y = f(x + k)$$

is a k-unit *leftward* shift of the graph of $y = f(x)$ and the graph of

$$y = f(x - k)$$

is a k-unit *rightward* shift of the graph of $y = f(x)$.

WARNING!! A common mistake is to think that the graph of $y = f(x + k)$ is a *rightward* shift of the graph of $y = f(x)$ when k is positive. This is incorrect; make sure you see why.

(c) The graph of $y = f(x+3) - 4$ is the result of shifting the graph of $y = f(x)$ downward 4 units and leftward 3 units. The leftward shift means the domain of h is 3 units "to the left" of the domain of f. The domain of f is $[-5, 3]$, so the domain of h is $[-8, 0]$. The downward shift means the range of h is 4 units "below" the range of f. The range of f is $[-3, 4]$, so the range of h is $[-7, 0]$.

□

In Problem 1.11, we learned the effect of adding a constant to the input or to the output of a function:

Important: When $k > 0$, the graph of $y = f(x) + k$ results from shifting the graph of $y = f(x)$ *vertically* upward by k units, while the graph of $y = f(x+k)$ results from shifting the graph of $y = f(x)$ *horizontally* to the left by k units. Similarly, the graph of $y = f(x) - k$ is a k-unit downward shift of the graph of $y = f(x)$ and the graph of $y = f(x - k)$ is a k-unit rightward shift.

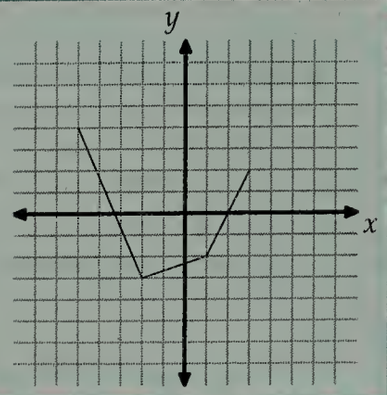
Problem 1.12: At right is the graph of $y = f(x)$.

(a) Find the graph of $y = -f(x)$.

(b) Find the graph of $y = f(-x)$.

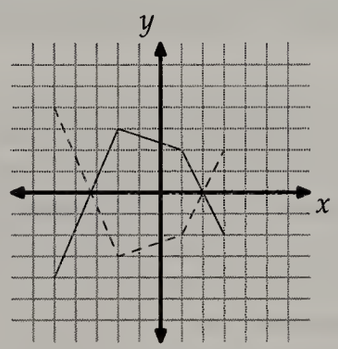
(c) Find the graph of $y = 2f(x)$.

(d) Find the graph of $y = f(2x)$.

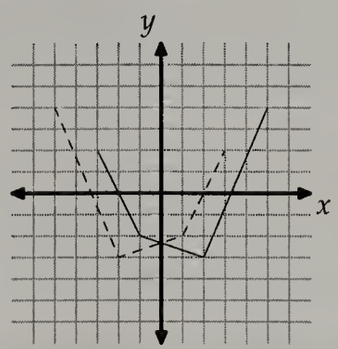


Solution for Problem 1.12:

(a) As before, we let $x = a$, where a is in the domain of f. Then, the point $(a, f(a))$ is on the graph of $y = f(x)$ and the point $(a, -f(a))$ is on the graph of $y = -f(x)$. So, the y-coordinate of each point on the graph of $y = -f(x)$ is the opposite of the y-coordinate of the corresponding point on the graph of $y = f(x)$. Therefore, we form the graph of $y = -f(x)$ by reflecting the graph of $y = f(x)$ over the x-axis. Below, the graph of $y = -f(x)$ is solid and the graph of $y = f(x)$ is dashed.

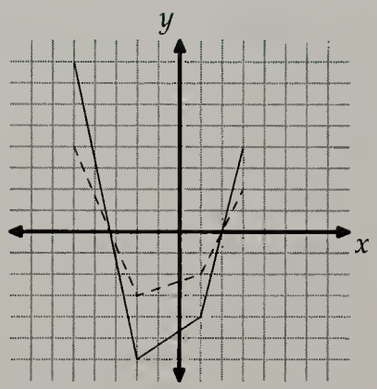


(b) We let $x = a$, where a is in the domain of f. Then, the point $(a, f(a))$ is on the graph of $y = f(x)$ and the point $(a, f(-a))$ is on the graph of $y = f(-x)$. We can find a point on the graph of $y = f(-x)$ with $f(a)$ as the y-coordinate by letting $x = -a$, which tells us that $(-a, f(a))$ is on the graph of $y = f(-x)$. Therefore, the graph of $y = f(-x)$ is the result of reflecting the graph of $y = f(x)$ over the y-axis. Below, the graph of $y = f(-x)$ is solid and the graph of $y = f(x)$ is dashed.

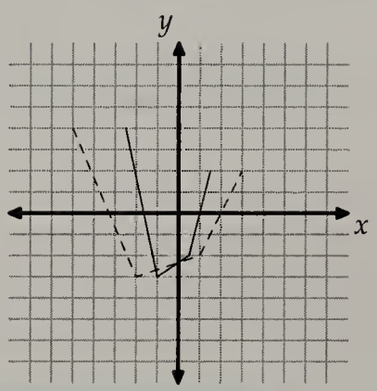


Important: The graph of $y = -f(x)$ is the result of reflecting the graph of $y = f(x)$ over the x-axis.
The graph of $y = f(-x)$ is the result of reflecting the graph of $y = f(x)$ over the y-axis.

(c) For each point $(a, f(a))$ on the graph of $y = f(x)$, the point $(a, 2f(a))$ is on the graph of $y = 2f(x)$. The latter point is twice as far from the x-axis as the former. Similarly, each point on the graph of $y = 2f(x)$ can be found by doubling the distance from the x-axis of the corresponding point on the graph of $y = f(x)$. Therefore, the graph of $y = 2f(x)$ can be found by scaling the graph of $y = f(x)$ vertically away from the x-axis by a factor of 2. At right, the graph of $y = 2f(x)$ is solid and the graph of $y = f(x)$ is dashed.



(d) For each point $(2a, f(2a))$ on the graph of $y = f(x)$, the point $(a, f(2a))$ is on the graph of $y = f(2x)$. The latter point is *half* as far from the y-axis as the former. Similarly, each point on the graph of $y = f(2x)$ can be found by halving the distance from the y-axis of the corresponding point on the graph of $y = f(x)$. Therefore, the graph of $y = f(2x)$ can be found by scaling the graph of $y = f(x)$ horizontally relative to the y-axis by a factor of $1/2$. At right, the graph of $y = f(2x)$ is solid and the graph of $y = f(x)$ is dashed.



□

Important: When $k > 0$, the graph of $y = kf(x)$ results from scaling the graph of $y = f(x)$ *vertically* relative to the x-axis by a factor of k, while the graph of $y = f(kx)$ results from scaling the graph of $y = f(x)$ *horizontally* relative to the y-axis by a factor of $1/k$.

WARNING!! The graph of $y = f(kx)$ is a horizontal scaling of the graph of $y = f(x)$ by a factor of $1/k$, not by k. So, for example, the graph of $y = f(3x)$ is the result of compressing the graph of $y = f(x)$ towards the y-axis, not stretching it away from the y-axis.

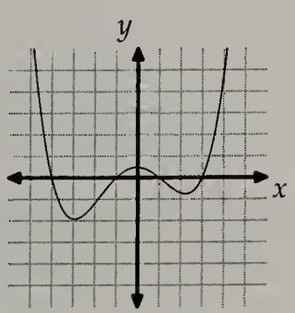
Exercises

1.2.1 Graph each of the following, and find the x-intercepts and y-intercepts of each graph.

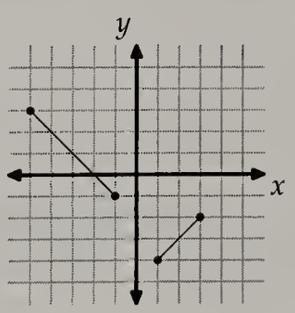
(a) $f(x) = 3x - 7$ (b) $f(x) = 2 - |x|$ (c) $f(x) = x^2 - 5x + 6$

1.2.2 Determine if each of the following graphs represents a function:

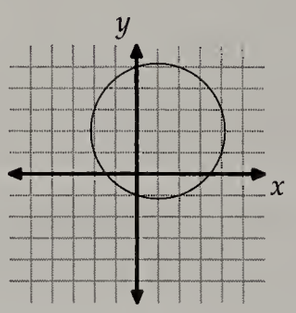
(a)



(b)



(c)



1.2.3 Suppose $f(9) = 2$. For each of the parts below, find a point that must be on the graph of the given equation.

(a) $y = f(x-3)+5$

(b) $y = 2f(x/4)$

(c) $y = 2f(3x-1)+7$

1.2.4★ Suppose $f(4) = 8$. Note that the point $(2,8)$ is on the graph of $y = f(2x)$ and that the point $(5,8)$ is on the graph of $y = f(2x-6)$. So, the point we have found on the graph of $y = f(2x-6)$ is 3 units to the right of the point we found on the graph of $y = f(2x)$. Which of the following is true: (a) we made a mistake; (b) it's a coincidence—not every point on $y = f(2x-6)$ is 3 to the right of a point on $y = f(2x)$; (c) the graph of $y = f(2x-6)$ is a 3-unit rightward shift of the graph of $y = f(2x)$. Explain your answer. **Hints:** 53

1.3 Composition

Sometimes we want to hook two (or more) machines together, taking the output from one machine and putting it into another. When we connect functions together like this, we are performing a **composition** of the functions. For example, the expression $f(g(x))$ means we put our input x into function g, and then take the output, $g(x)$, and put that into function f. A composition of functions is also sometimes indicated with the symbol $\circ$. For example, when we write $h = f \circ g$, we define the function h such that $h(x) = f(g(x))$.

Problems

Problem 1.13: Suppose $f(x) = 3x+7$ and $g(x) = 2\sqrt{x-3}$.

(a) Find $f(g(3))$ and $g(f(3))$.

(b) Find a if $f(g(a)) = 25$.

(c) Is $g(f(-7))$ defined? Why or why not?

(d) Suppose that $h = g \circ f$. What is the domain of h?

Problem 1.14:

(a) Let $h(x) = f(g(x))$, where f and g are functions, and suppose $h(x)$ is defined for all values of x in the domain of g. Explain why the range of g must be part of the domain of f.

(b) Again, let $h(x) = f(g(x))$, where f and g are functions. Suppose every value in the range of g is in the domain of f. Must the domain of h be the same as the domain of g?

Problem 1.15: Let $f(a) = a-2$ and $F(a,b) = b^2 + a$. Find $F(3, f(4))$. *(Source: AHSME)*

Problem 1.16:

(a) Let $f(x) = 3x-1$. Find $f(f(x))$ and $f(f(f(x)))$.

(b) Suppose $g(x) = ax+b$, where a and b are real constants. Find all possible pairs (a,b) such that we have $g(g(x)) = 9x+28$.

Problem 1.17: Let $f(x) = \sqrt{x}$ and $g(x) = x^2$. Is it true that $f(g(x))$ and $g(f(x))$ are the same function? Why or why not?

Problem 1.13: Suppose $f(x) = 3x + 7$ and $g(x) = 2\sqrt{x-3}$.

(a) Find $f(g(3))$ and $g(f(3))$.

(b) Find a if $f(g(a)) = 25$.

(c) Is $g(f(-7))$ defined? Why or why not?

(d) Suppose that $h = g \circ f$. What is the domain of h?

Solution for Problem 1.13:

(a) We have $g(3) = 2\sqrt{3-3} = 0$, so $f(g(3)) = f(0) = 3(0) + 7 = 7$. Similarly, we have $f(3) = 3(3) + 7 = 16$, so $g(f(3)) = g(16) = 2\sqrt{16-3} = 2\sqrt{13}$. Notice that $f(g(3)) \neq g(f(3))$.

WARNING!! If $f(x)$ and $g(x)$ are functions, then the functions $f(g(x))$ and $g(f(x))$ are *not necessarily the same function.*

(b) We have

$$f(g(a)) = f(2\sqrt{a-3}) = 3(2\sqrt{a-3}) + 7 = 6\sqrt{a-3} + 7,$$

so $f(g(a)) = 25$ means

$$6\sqrt{a-3} + 7 = 25.$$

Subtracting 7 from both sides and then dividing both sides by 6 gives $\sqrt{a-3} = 3$. Squaring both sides gives us $a - 3 = 9$, so $a = 12$. Checking our answer, we find $f(g(12)) = f(6) = 25$, as expected.

(c) We have $f(-7) = 3(-7) + 7 = -14$, so $g(f(-7)) = g(-14)$. However, the value -14 is not in the domain of g. Since $g(-14)$ is not defined, $g(f(-7))$ is not defined.

(d) Since $h(x) = g(f(x))$, in order for x to be in the domain of h, both $f(x)$ and $g(f(x))$ must be defined. The domain of f is all real numbers; however, as we saw in the previous part, $g(f(x))$ is not defined for all real values of x.

In order for $g(f(x))$ to be defined, the value of $f(x)$ must be in the domain of g. The domain of g is $[3, +\infty)$. Therefore, in order for $g(f(x))$ to be defined, we must have $f(x) \geq 3$. So, we must have $3x + 7 \geq 3$, which gives us $x \geq -4/3$. For all values of x such that $x \geq -4/3$, the expression $g(f(x))$ is defined. For any value of x less than $-4/3$, the expression $g(f(x))$ is not defined. Therefore, the domain of h is $[-4/3, +\infty)$.

□

Problem 1.14:

(a) Let $h(x) = f(g(x))$, where f and g are functions, and suppose $h(x)$ is defined for all values of x in the domain of g. Explain why the range of g must be part of the domain of f.

(b) Again, let $h(x) = f(g(x))$, where f and g are functions. Suppose every value in the range of g is in the domain of f. Must the domain of h be the same as the domain of g?

Solution for Problem 1.14:

(a) Since $h(x) = f(g(x))$ and $h(x)$ is defined for all values of x in the domain of g, the expression $f(g(x))$ must be defined for all values of x in the domain of g. Because $f(g(x))$ is defined, the value of $g(x)$ is in the domain of f. So, every number in the range of g must be in the domain of f.

(b) The expression $f(g(x))$ is not defined if $g(x)$ is not defined. Therefore, all values in the domain of h must be in the domain of g. Conversely, if x is in the domain of g, then $g(x)$ is defined. Furthermore, we are given that every value in the range of g is in the domain of f, so $f(g(x))$ is defined for all x in the domain of g. Therefore, every value in the domain of g is in the domain of h. Since every value in the domain of g is in the domain of h, and vice versa, we conclude that the domain of h is the same as the domain of g. □

Problem 1.15: If $f(a) = a - 2$ and $F(a, b) = b^2 + a$, find the value of $F(3, f(4))$. *(Source: AHSME)*

Solution for Problem 1.15: We note that $f(4) = 4 - 2 = 2$, so

$$F(3, f(4)) = F(3, 2) = 2^2 + 3 = 7.$$

□

Although the expression $F(3, f(4))$ looks a little intimidating, the process of evaluating it is just simple plug-and-chug arithmetic.

Concept: Don't be intimidated by notation. Often the notation is a fancy way of representing very simple ideas.

We'll see the value of notation as we move to more complicated problems throughout this book. However, we'll usually introduce notation with simple problems like Problem 1.15 so you can get used to the notation before using it on harder problems.

Problem 1.16:

(a) Let $f(x) = 3x - 1$. Find $f(f(x))$ and $f(f(f(x)))$.

(b) Suppose $g(x) = ax + b$, where a and b are real constants. Find all possible pairs (a, b) such that we have $g(g(x)) = 9x + 28$.

Solution for Problem 1.16:

(a) We have

$$f(f(x)) = f(3x - 1) = 3(3x - 1) - 1 = 9x - 4.$$

We can now use this result to find $f(f(f(x)))$ by inputting $f(f(x))$, which equals $9x - 4$, into $f(x)$:

$$f(f(f(x))) = f(9x - 4) = 3(9x - 4) - 1 = 27x - 13.$$

(b) In order to find a and b, we first need an expression for $g(g(x))$. We have

$$g(g(x)) = g(ax + b) = a(ax + b) + b = a^2x + ab + b.$$

So, our equation $g(g(x)) = 9x + 28$ is now $a^2x + ab + b = 9x + 28$. In order for this equation to be true for all values of x, the coefficient of x on both sides must be the same and the constant terms on both sides must be the same. Therefore, we have the system of equations

$$a^2 = 9,$$
$$ab + b = 28.$$

From $a^2 = 9$, we have $a = \pm 3$. Letting $a = 3$ in the second equation gives $b = 7$ and letting $a = -3$ in the second equation gives $b = -14$. So, our two possible pairs are $(a, b) = (3, 7)$ and $(a, b) = (-3, -14)$.

□

We have a special notation for repeatedly feeding the output of a function back into the function itself. In this text, we will often write $f^2(x)$ to mean $f(f(x))$. Likewise, $f^n(x)$ usually refers to applying the function $f(x)$ exactly n times; for example, $f^5(x) = f(f(f(f(f(x)))))$.

WARNING!! Some sources use $f^2(x)$ to mean $f(x) \cdot f(x)$ instead of $f(f(x))$. Likewise, $f^n(x)$ is sometimes used to denote $[f(x)]^n$. You'll often have to determine by context whether $f^n(x)$ refers to a function applied n times, or to a function raised to the n^{th} power.

Problem 1.17: Let $f(x) = \sqrt{x}$ and $g(x) = x^2$. Is it true that $f(g(x))$ and $g(f(x))$ are the same function? Why or why not?

Solution for Problem 1.17: At first glance, we might be tempted to reason as follows:

Bogus Solution: We have $f(g(x)) = \sqrt{x^2} = x$ and $g(f(x)) = (\sqrt{x})^2 = x$, so $f(g(x))$ and $g(f(x))$ are the same.

This Bogus Solution contains two significant errors. First, the domain of f is all nonnegative numbers, so the domain of $g(f(x))$ is all nonnegative numbers. However, because the domain of g is all reals while the range of g is all nonnegative numbers, the expression $f(g(x))$ is defined for all real x. Therefore, the domain of $f(g(x))$ is all real numbers. So, the two functions $g(f(x))$ and $f(g(x))$ have different domains, which means that the functions are not the same. For example, we have $f(g(-1)) = f(1) = 1$, but $g(f(-1))$ is undefined, because $f(-1)$ is undefined.

The other error occurs when we write $\sqrt{x^2} = x$. This statement is only true if $x \geq 0$. If $x < 0$, then we have $\sqrt{x^2} = -x$. So, it is not true that $f(g(x)) = x$ for all x. If $x \geq 0$, then $f(g(x)) = x$, and if $x < 0$, then $f(g(x)) = -x$. We can capture both of these by writing $f(g(x)) = |x|$. □

Exercises

1.3.1 Check the answer to part (b) of Problem 1.16 by evaluating $g(g(x))$ for $g(x) = 3x + 7$ and for $g(x) = -3x - 14$.

1.3.2 Let $f(x) = x^2 - 2x$ and $g(x) = \sqrt{1 - x}$.

(a) Find $f(g(x))$.

(b) Find the domain of $f(g(x))$.

(c) Is $g(f(-2))$ defined? Why or why not?

1.3.3 Are both $f(f^2(x))$ and $f^2(f(x))$ the same as $f^3(x)$?

1.3.4 Let a be a constant, and let $f(x) = ax$. Find an expression for $f^n(x)$ in terms of a and n, where n is a positive integer.

1.3.5 Suppose h, f, and g are functions such that $h(x) = f(g(x))$. If $(3, 0)$ is the only x-intercept of the graph of g, then can we determine the x-intercepts or y-intercepts of h?

1.4 Inverse Functions

You're probably very familiar with using Control-Z when typing on a computer: that's the "undo" command for many applications. Being able to "undo" the work of a function can be extremely useful. The machine that "undoes" the work of a function f is called the **inverse function** of f.

Definition: Function g is the **inverse** of function f if

$$g(f(x)) = x \text{ for all values of } x \text{ in the domain of } f, \text{ and}$$
$$f(g(x)) = x \text{ for all values of } x \text{ in the domain of } g.$$

From our definition, we see that if g is the inverse function of f, then f is the inverse function of g. Moreover, the range of g is the domain of f, and vice versa.

As we will see, not every function has an inverse. However, if a function has an inverse, then the inverse is unique, as our definition implies. We can see this by noting that if g and h are both inverses of f, then we have $g(f(x)) = x$ and $h(f(x)) = x$, so $g(f(x)) = h(f(x))$ for all x in the domain of f. Letting $y = f(x)$, we have $g(y) = h(y)$ for all y in the domain of g and h, so g and h are the same function.

We have a special notation to denote the inverse of a function. We often write the inverse of f as f^{-1}.

WARNING!! When working with functions, f^{-1} does not mean $\frac{1}{f}$! The expression f^{-1} is a special notation that denotes the inverse of the function f.

Problems

Problem 1.18: In this problem we find the inverse of the function $f(x) = 3x - 5$.

(a) Let g be the inverse of f, so that we have $f(g(x)) = x$ and $g(f(x)) = x$. Which of these equations is easier to use to find $g(x)$?

(b) Use the definition of f to solve the equation $f(y) = x$ for y.

(c) What is $g(x)$?

(d) Check your answer by confirming that $g(f(x)) = x$.

Problem 1.19: Let $f(x) = \dfrac{-5 - x}{3 - 2x}$.

(a) Find $f^{-1}(7)$, *without finding* $f^{-1}(x)$.

(b) Find $f^{-1}(x)$.

Problem 1.20:

(a) Does the function $g(x) = 2x^2 - 9$ have an inverse? If it does, find it. If it doesn't, explain why it doesn't.

(b) Does the function $h(x) = 3\sqrt{2x + 1}$ have an inverse? If it does, find it. If it doesn't, explain why it doesn't. (Be careful on this part!)

Problem 1.21:

(a) Explain how to determine from the graph of a function whether or not the function has an inverse.

(b) How are the graph of a function and the graph of its inverse related?

Problem 1.22: Suppose functions f and g both have inverses. Show that if $h = f \circ g$, then the inverse of h is $g^{-1} \circ f^{-1}$.

Problem 1.18: Find the inverse of the function $f(x) = 3x - 5$.

Solution for Problem 1.18: Suppose g is the inverse of f. Then, we must have

$$g(f(x)) = f(g(x)) = x.$$

The equation $f(g(x)) = x$ is more helpful, because we know $f(x)$. So, we can put $g(x)$ into $f(x)$, which gives us $f(g(x)) = 3g(x) - 5$. Therefore, we must have $3g(x) - 5 = x$.

Important: Just as we can sometimes solve an equation for a variable by isolating the variable, we can sometimes solve an equation for a function by isolating that function.

We isolate $g(x)$ by adding 5 to both sides, then dividing by 3. This gives us $g(x) = \frac{x+5}{3}$. We can check our work by confirming that $f(g(x)) = g(f(x)) = x$. Therefore, we have $f^{-1}(x) = \frac{x+5}{3}$.□

Note that we found the inverse of f by solving the equation $f(g(x)) = x$ for $g(x)$. Make sure you see why this produces a function g that is the inverse of f.

Important: If f has an inverse g, then we can often find that inverse by solving the equation $f(g(x)) = x$ for $g(x)$.

If you find the idea of "solving for a function" confusing, you can instead let $y = g(x)$ and solve for y. For example, when $f(x) = 3x - 5$, we can solve for y in $f(y) = x$ to find the inverse of f. Since $f(y) = x$, we have $3y - 5 = x$. Solving for y gives $y = \frac{x+5}{3}$, so the inverse of f is $f^{-1}(x) = \frac{x+5}{3}$.

Problem 1.19: Let $f(x) = \dfrac{-5-x}{3-2x}$.

(a) Find $f^{-1}(7)$, *without finding* $f^{-1}(x)$.

(b) Find $f^{-1}(x)$.

Solution for Problem 1.19:

(a) Suppose $y = f^{-1}(7)$. Because f^{-1} is the inverse of f, we have $f(f^{-1}(x)) = x$ for all values of x. Specifically, we have $f(f^{-1}(7)) = 7$, so because $y = f^{-1}(7)$, we have $f(y) = 7$. We also could have deduced $f(y) = 7$ by noting that $y = f^{-1}(7)$ means that the y is the input to f that produces 7 as an output.

Important: If function f has an inverse, then $f(a) = b$ means that $f^{-1}(b) = a$, and vice versa.

Since $f(y) = 7$, we have

$$\frac{-5-y}{3-2y} = 7.$$

Multiplying both sides by $3 - 2y$ gives $-5 - y = 7(3 - 2y)$. Solving this equation gives us $y = 2$. Therefore, $f^{-1}(7) = 2$.

(b) As we just learned, we can find the inverse of f by solving the equation $f(y) = x$ for y in terms of x. Using our definition of f, the equation $f(y) = x$ becomes

$$\frac{-5-y}{3-2y} = x.$$

Multiplying both sides by $3 - 2y$ gives $-5 - y = x(3 - 2y)$. Isolating the terms with y on the left side gives $2xy - y = 3x + 5$, so $y(2x - 1) = 3x + 5$. Dividing by $2x - 1$ gives

$$f^{-1}(x) = y = \frac{3x+5}{2x-1}.$$

With this, we can check our answer to the first part by directly evaluating $f^{-1}(7) = 2$.

□

We say that a function is **invertible** if it has an inverse. As you might have guessed, not all functions are invertible.

Problem 1.20:
(a) Does the function $g(x) = 2x^2 - 9$ have an inverse? If it does, find it. If it doesn't, explain why it doesn't.
(b) Does the function $h(x) = 3\sqrt{2x+1}$ have an inverse? If it does, find it. If it doesn't, explain why it doesn't.

Solution for Problem 1.20:

(a) Both $g(1)$ and $g(-1)$ equal -7. If g has an inverse, should this inverse return 1 or -1 when -7 is input? We can't tell! Therefore, g is not reversible; it does not have an inverse.

Important: If a function gives the same output for two different inputs, then the function does not have an inverse.

On the other hand, if a function never gives the same output for two different inputs, then the function is invertible—we can tell from the output of the function what input produced that output.

Important: If there are no two different inputs to a function that produce the same output, then the function has an inverse.

(b) We start by looking for two different inputs to h that give the same output. After inputting a few values to h, we notice that as we input higher and higher values of x, we get higher and higher values of $h(x)$. This makes us think that h might indeed have an inverse, since it seems impossible to find two different inputs that give the same output. So, we try to find the inverse of h by solving the equation $h(y) = x$ for y. This gives us the equation

$$3\sqrt{2y+1} = x.$$

Squaring both sides gives us $9(2y+1) = x^2$. Solving for y gives

$$y = \frac{x^2}{18} - \frac{1}{2}.$$

This isn't the end of the story, however. Recall our definition of inverse functions. Functions f and h are inverse functions of each other if

$$h(f(x)) = x \text{ for all values of } x \text{ in the domain of } f, \text{ and}$$
$$f(h(x)) = x \text{ for all values of } x \text{ in the domain of } h.$$

We let $f(x) = \frac{x^2}{18} - \frac{1}{2}$. Our work above suggests that this is the inverse of h. However, if f is the inverse of h, then h is the inverse of f. So, we should have $h(f(x)) = x$ for all x in the domain of f. Unfortunately, we have $f(-3) = 0$, but $h(f(-3)) = h(0) = 3$, so we don't have $h(f(x)) = x$ when $x = -3$. The problem here is that the function $f(x) = \frac{x^2}{18} - \frac{1}{2}$ *does not have an inverse.* We can see this quickly by noting that $f(3) = f(-3) = 0$. So, where did we go wrong?

Let's evaluate $h(f(x))$ and see if it really does equal x. Maybe this will shed some light on where we went wrong. We have

$$h(f(x)) = 3\sqrt{2f(x)+1} = 3\sqrt{\frac{x^2}{9} - 1 + 1} = 3\sqrt{\frac{x^2}{9}} = \sqrt{x^2}.$$

Here, we have to be careful.

WARNING!! We only have $\sqrt{x^2} = x$ if x is nonnegative.

Now, we've found our problem. We only have $h(f(x)) = x$ if $\sqrt{x^2} = x$. This only occurs if x is nonnegative. So, if we restrict the domain of f to nonnegative numbers, then we have $h(f(x)) = x$ for all values of x in the domain of f. We can quickly confirm that $f(h(x)) = x$ for all x in the domain of h as well. So, the inverse of h is $h^{-1}(x) = \frac{x^2}{18} - \frac{1}{2}$, where the function h^{-1} is only defined for nonnegative values of x.

□

Problem 1.21:

(a) Explain how to determine from the graph of a function whether or not the function has an inverse.

(b) How are the graph of a function and the graph of its inverse related?

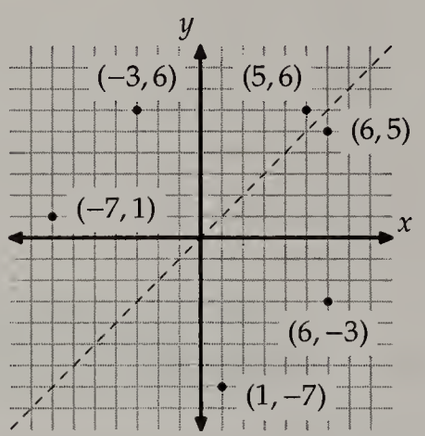
Solution for Problem 1.21:

(a) A function has an inverse if and only if we can always tell from the output of the function what the input was. If f has an inverse, then when we graph $y = f(x)$, there can be no two points on the graph with the same y-coordinate. If there were, then we would have two different inputs (the x-coordinates of points on the graph) that provide the same output (the common y-coordinate) for the function, so it couldn't have an inverse. "No two points on the graph have the same y-coordinate" means the same thing as "No horizontal line intersects the graph in more than one point." Therefore, we have a **horizontal line test** for whether or not a function has an inverse:

Important: A function has an inverse if and only if there does not exist a horizontal line that passes through more than one point on the graph of the function.

(b) Suppose the point (a, b) is on the graph of $y = f(x)$. Then, we have $f(a) = b$, so $f^{-1}(b) = a$. Because $f^{-1}(b) = a$, the point (b, a) is on the graph of $y = f^{-1}(x)$. Therefore, if (a, b) is on the graph of a function, then (b, a) is on the graph of its inverse.

The fact that the points on the graph of $y = f^{-1}(x)$ can be found by reversing the coordinates of the points on the graph of $y = f(x)$ has a geometric interpretation. We just learned that if (a, b) is on the graph of $y = f(x)$, then (b, a) is on the graph of $y = f^{-1}(x)$. In the graph at right, we have plotted the points $(5, 6)$, $(6, -3)$, and $(-7, 1)$. We have also plotted the points that result when the coordinates of these three points are reversed. We see that the new points are the result of reflecting the old points over the line $y = x$. The graph of $y = x$ is the dashed line in the diagram at right.



In general, reversing the coordinates of a point has the same effect as reflecting the point over the graph of $y = x$. To prove this, we let point P be (a, b) and Q be (b, a). The slope of $\overline{PQ}$ is -1, so $\overline{PQ}$ is perpendicular to the graph of $y = x$. Furthermore, the midpoint of $\overline{PQ}$ is $(\frac{a+b}{2}, \frac{a+b}{2})$, which is on the graph of $y = x$. So, the graph of $y = x$ is the perpendicular bisector of $\overline{PQ}$, which means that P and Q are images of each other upon reflection over the graph of $y = x$. So, each point on the graph of $y = f^{-1}(x)$ is the mirror image of a point on the graph of $y = f(x)$ when reflected over the line $y = x$ (and vice versa). Therefore, the entire graph of $y = f^{-1}(x)$ is the mirror image of the graph of $y = f(x)$ when reflected over the line $x = y$. □

Important: If the function f has an inverse f^{-1}, then the graph of $y = f^{-1}(x)$ is the reflection of the graph of $y = f(x)$ over the line $x = y$.

Problem 1.22: Suppose functions f and g both have inverses. Show that if $h = f \circ g$, then the inverse of h is $g^{-1} \circ f^{-1}$.

Solution for Problem 1.22: Let $t = g^{-1} \circ f^{-1}$. To show that t is the inverse of h, we must show that $h(t(x)) = x$ and $t(h(x)) = x$. We'll start with $h(t(x))$. Using our definitions of h and t, we have

$$h(t(x)) = f(g(g^{-1}(f^{-1}(x)))).$$

That looks pretty scary! But seeing both g and g^{-1} right next to each other reminds us that $g(g^{-1}(x)) = x$. In other words, the output of $g \circ g^{-1}$ is always the same as the input. So, when we put $f^{-1}(x)$ into $g \circ g^{-1}$, we get $f^{-1}(x)$ out, which means

$$g(g^{-1}(f^{-1}(x))) = f^{-1}(x).$$

So, using this in our expression for $h(t(x))$ above, we have

$$h(t(x)) = f(g(g^{-1}(f^{-1}(x)))) = f(f^{-1}(x)).$$

Because f^{-1} is the inverse of f, we have $f(f^{-1}(x)) = x$, so $h(t(x)) = x$.

We can do essentially the same thing for $t(h(x))$. We start with

$$t(h(x)) = g^{-1}(f^{-1}(f(g(x)))).$$

Because f is the inverse of f^{-1}, we have $f^{-1}(f(g(x))) = g(x)$, so $t(h(x)) = g^{-1}(f^{-1}(f(g(x)))) = g^{-1}(g(x)) = x$, as desired. Because $h(t(x)) = t(h(x)) = x$ when $h = f \circ g$ and $t = g^{-1} \circ f^{-1}$, the function $g^{-1} \circ f^{-1}$ is the inverse of $f \circ g$. □

Exercises

1.4.1 If $f(x) = \frac{x-1}{x-2}$, what is the value of $f^{-1}(3)$?

1.4.2 Determine whether or not each of the following functions has an inverse. If the function has an inverse, find it.

(a) $f(x) = 2x - 7$

(b) $f(x) = \dfrac{1}{2x+3}$

(c) $f(x) = \sqrt{2-x}$

(d) $f(x) = |5 - x|$

(e) $f(x) = \sqrt{4-x} + \sqrt{x-2}$

(f)★ $f(x) = x^2 - 6x + 3$, where $x \geq 4$

1.4.3 Describe all ordered pairs (a, b) such that the inverse of $f(x) = ax + b$ satisfies $f(x) = f^{-1}(x)$ for all x.

1.4.4 Let $g(x) = \frac{ax+b}{cx+d}$, where a, b, c, and d are positive and $ad \neq bc$. Which of $\frac{a}{b}$, $\frac{a}{c}$, and $\frac{a}{d}$ *cannot* be in the domain of g^{-1}? **Hints:** 146

1.5 Summary

A **function** gives a single output for each input. The **domain** of a function consists of all the values we are able to input to the function and get an output. Meanwhile, the **range** of the function consists of all the values that can come out of the function.

When we graph the function f, we graph the equation $y = f(x)$ on the Cartesian plane. Therefore, if the point (a, b) is on the graph of $y = f(x)$, then $f(a) = b$. Conversely, if $f(a) = b$, then (a, b) is on the graph of $y = f(x)$.

Important: A curve on the Cartesian plane is the graph of some function if and only if every vertical line passes through no more than one point on the curve. We call this the **vertical line test**.

Important: When $k > 0$, the graph of $y = f(x) + k$ results from shifting the graph of $y = f(x)$ *vertically* upward by k units, while the graph of $y = f(x+k)$ results from shifting the graph of $y = f(x)$ *horizontally* to the left by k units. Similarly, the graph of $y = f(x) - k$ is a k-unit downward shift of the graph of $y = f(x)$ and the graph of $y = f(x - k)$ is a k-unit rightward shift.

Important: The graph of $y = -f(x)$ is the result of reflecting the graph of $y = f(x)$ over the x-axis. The graph of $y = f(-x)$ is the result of reflecting the graph of $y = f(x)$ over the y-axis.

Important: When $k > 0$, the graph of $y = kf(x)$ results from scaling the graph of $y = f(x)$ *vertically* by a factor of k, while the graph of $y = f(kx)$ results from scaling the graph of $y = f(x)$ *horizontally* by a factor of $1/k$.

When we perform a **composition** of two functions, we place the output from the first function into the second. For example, both $f(g(x))$ and $(f \circ g)(x)$ indicate that we take the output $g(x)$ and input this result to f. We have a special notation for when we feed the output of a function back into the function itself. In this text, we write $f^2(x)$ to mean $f(f(x))$. Likewise, $f^n(x)$ refers to applying the function $f(x)$ exactly n times; for example, $f^5(x) = f(f(f(f(f(x)))))$.

Finally, we have a special name for a pair of functions such that each function "undoes" the action of the other:

Definition: If f and g are functions such that

$$g(f(x)) = x \text{ for all values of } x \text{ in the domain of } f, \text{ and}$$
$$f(g(x)) = x \text{ for all values of } x \text{ in the domain of } g,$$

then f is the **inverse** of g, and g is the **inverse** of f.

Not every function has an inverse. If a function gives the same output for two different inputs, then the function does not have an inverse. If there are no two different inputs to a function that produce the same output, then the function has an inverse. If f has an inverse g, then we can often find that inverse by solving the equation $f(g(x)) = x$ for $g(x)$.

We usually denote the inverse of the function f as f^{-1}. If f has an inverse, then the graph of f^{-1} is the reflection of the graph of f over the line $y = x$.

Important: A function has an inverse if and only if there does not exist a horizontal line that passes through more than one point on the graph of the function.

Things To Watch Out For!

- When evaluating a function that accepts multiple inputs, make sure you assign the values to the dummy variables in the correct order.
- We must take into account any restrictions on the domain of a function when finding the range of the function.
- Sometimes there is more than one constraint on the domain of a function.

- A common mistake is to think that the graph of $y = f(x+k)$ is a *rightward* shift of the graph of $y = f(x)$ when k is positive. This is incorrect; make sure you see why.
- The graph of $y = f(kx)$ is a horizontal scaling of the graph of $y = f(x)$ by a factor of $1/k$, not by k. So, the graph of $y = f(3x)$ is the result of compressing the graph of $y = f(x)$ towards the y-axis, not stretching it away from the y-axis.
- Some sources use $f^2(x)$ to mean $f(x) \cdot f(x)$ instead of $f(f(x))$. Likewise, $f^n(x)$ is sometimes used to denote $[f(x)]^n$. You'll often have to determine by context whether $f^n(x)$ refers to a function applied n times, or to a function raised to the n^{th} power.
- Don't be intimidated by notation. Often the notation is a fancy way of representing very simple ideas.
- When working with functions, f^{-1} does not mean $\frac{1}{f}$! The expression f^{-1} is a special notation that denotes the inverse of the function f.

Problem Solving Strategies

Concepts:

- We can often check our answers in algebra problems by plugging the solutions we find back into the original question.
- Keep your eye on the ball. Sometimes we don't need to evaluate every variable in a problem in order to solve the problem.

REVIEW PROBLEMS

1.23 Find the domain and range of each of the following functions:

(a) $A(x) = 4x^2 + 1$

(b) $o(x) = 3 + \sqrt{16 - (x-3)^2}$

(c) $P(x) = \dfrac{1}{3 + \sqrt{x+1}}$

(d) $S(x) = \dfrac{12x - 9}{6 - 9x}$

1.24 Find the domain of each of the following functions:

(a) $f(x) = \dfrac{1}{\sqrt{2x-5}} + \sqrt{9 - 3x}$

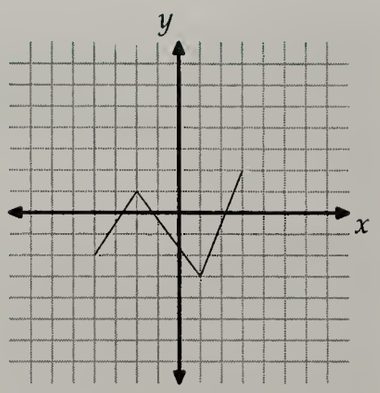
(b) $f(x) = |\sqrt{x} - 2| + |\sqrt{x-2}|$

(c) $g(x) = \sqrt{|x| - 2} + \sqrt{|x-3|}$

1.25 Let $f(x) = \dfrac{4x}{x+2}$ and $g(x) = \dfrac{2x}{x+4}$. Find $f(g(x))$.

1.26 Let $f(x) = ax^2 + bx + c$ and $g(x) = ax^2 - bx + c$. If $f(1) = g(1) + 2$ and $f(2) = 2$, find $g(2)$.

1.27 At right is the graph of $y = f(x)$. Graph each of the following:



(a) $y = f(x - 3)$

(b) $y = 2f(x) + 1$

(c) $y = f(-x)$

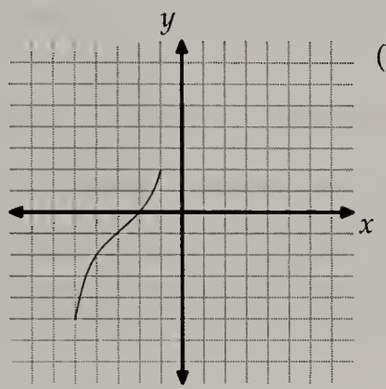
(d) $y = f(2 - x)$

(e)★ $y = \frac{1}{2}f(2x - 1) + 3$ **Hints:** 249

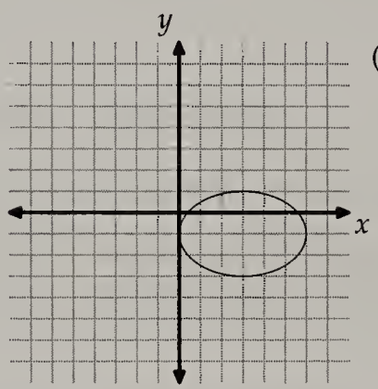
1.28 Suppose that $f(f(x)) = g(f(x))$ for all real x. Must f and g be the same function?

1.29 For each of the following three graphs, determine if the graph could represent a function. If the graph can represent a function, then determine if the function has an inverse. If it does have an inverse, then graph the inverse.

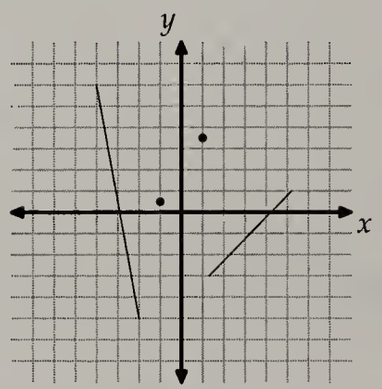
(a)



(b)



(c)



1.30 Let $f(x, y, z) = \dfrac{x - z}{y - z}$. If $f(a, b, c) = 1$, what is the value of $f(a, c, b)$?

1.31 Suppose that $f(x) = -x^2 + bx + c$ and $g(x) = dx + e$, where b, c, d, and e are real constants. Is it possible that $f(g(x)) = x^2$? Why or why not?

1.32 Must the inverse of a linear function $f(x) = ax + b$, where $a \neq 0$, be a linear function?

1.33 Suppose that a function $f(x)$ is defined for all real x. How can you use the graph of $f(x)$ to produce the graph of $y = |f(x)|$?

1.34 Let $f(t) = \dfrac{t}{1 - t}$, where $t \neq 1$. If $y = f(x)$, then x can be expressed as which of the following:

$$f(1/y),\ -f(y),\ -f(-y),\ f(-y),\ f(y)?$$

(Source: AHSME)

1.35 The function f defined by $f(x) = \dfrac{cx}{2x + 3}$ satisfies $f(f(x)) = x$ for all real numbers x except $-3/2$. Find c.
(Source: AHSME)

1.36 Suppose that $f(x)$ has an inverse, $f^{-1}(x)$. Must $f^{-1}(2x)$ be the inverse of $f(2x)$? If so, prove it. If not, find an expression that is the inverse of $f(2x)$.

1.37 The function $f(x) = x^2$ is not invertible. However, if we let $g(x) = \sqrt{x}$, then $f(g(x)) = (\sqrt{x})^2 = x$. So, why isn't f invertible?

Challenge Problems

1.38 Describe all constants a and b such that $f(x) = \dfrac{2x+a}{bx-2}$ and $f(x) = f^{-1}(x)$ for all x in the domain of f.

1.39 Suppose that a function $f(x)$ has domain $(-1, 1)$. Find the domains of the following functions:

(a) $f(x+1)$ (b) $f(1/x)$ (c) $f(\sqrt{x})$ (d)★ $f\left(\frac{x+1}{x-1}\right)$

1.40 Find the domain and range of the function $f(x) = \sqrt{2-x-x^2}$. **Hints:** 47

1.41 Let $f(x^2+1) = x^4 + 5x^2 + 3$. What is $f(x^2-1)$? *(Source: AMC 12)* **Hints:** 184

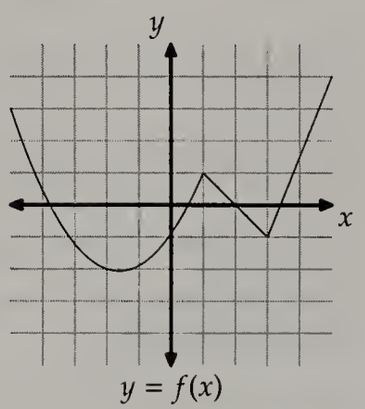
1.42 Suppose that a function $f(x)$ is defined for all real x. How can we obtain the graph of $f(|x|)$ from the graph of $f(x)$? **Hints:** 271

1.43 If $f(f(x)) = x$ for all x for which $f(x)$ is defined and $f(0) = 2003$, find all roots of the equation $f(x) = 0$.

1.44 For which constants a, b, c, and d does the function $f(x) = \dfrac{ax+b}{cx+d}$ have an inverse?

1.45 Suppose that $g(x) = f(x)f(|x|)$, where $f(x)$ is a function with $\mathbb{R}$ as its domain. Must all x-intercepts of the graph of f be x-intercepts of the graph of g? Must all x-intercepts of the graph of g be x-intercepts of the graph of f?

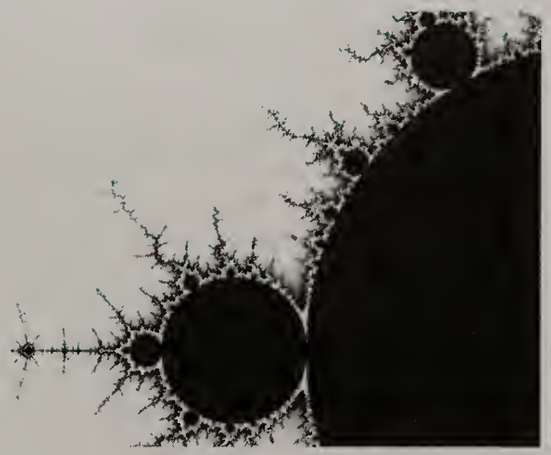
1.46 At right is the graph of $y = f(x)$. Draw the graph of $y = \dfrac{3}{2}f(2x-2) - 1$.



1.47★ The function f defined by

$$f(x) = \frac{ax+b}{cx+d},$$

where a, b, c, and d are nonzero real numbers, has the properties $f(19) = 19$, $f(97) = 97$, and $f(f(x)) = x$ for all values of x except $-d/c$. Find the unique number that is not in the range of f. *(Source: AIME)* **Hints:** 185



There is nothing more difficult to take in hand, more perilous to conduct or more uncertain in its success than to take the lead in the introduction of a new order of things. – Nicoolo Machiavelli

CHAPTER 2

Introduction to Trigonometric Functions

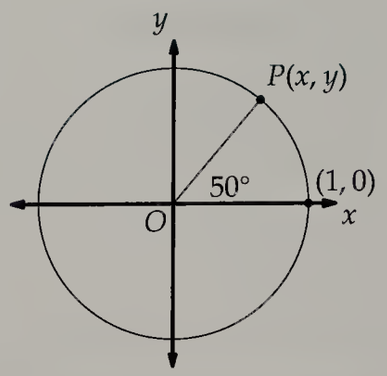
Trigonometry is the study of triangles, but when most people refer to "trigonometry," they speak of a special set of functions that are often useful when working with triangles. You may be familiar with some of these functions, such as **sine**, **cosine**, and **tangent**, from applying them to the acute angles of right triangles. In this text, we go well beyond right triangles, starting with a more complete definition of these **trigonometric functions**.

Sidenote: Throughout this text, you'll see some funny symbols used to represent angle measures. These funny symbols are lowercase Greek letters, which are often used to denote angle measures. Perhaps mathematicians use Greek letters this way because the Greeks contributed so much to geometry, or perhaps because the letters are so fun to write! Here's the lowercase Greek alphabet:

Symbol	Name	Symbol	Name
α	alpha	ν	nu
β	beta	ξ	xi
γ	gamma	o	omicron
ϵ	epsilon	ρ	rho
ζ	zeta	σ	sigma
η	eta	τ	tau
θ	theta	υ	upsilon
ι	iota	ϕ	phi
κ	kappa	χ	chi
λ	lambda	ψ	psi
μ	mu	ω	omega

2.1 The Unit Circle

We find the cosine, sine, and tangent of an angle θ by using a circle with radius 1 centered at the origin of the Cartesian plane. We call this circle the **unit circle**. We start by finding the point that results when the point $(1, 0)$ is rotated counterclockwise by θ about the origin. We call the resulting point the **terminal point** of θ. For example, point P in the diagram is the terminal point of $50°$, since it is the point that results when $(1, 0)$ is rotated $50°$ counterclockwise about the origin. Since the terminal point of an angle is the result of rotating about the origin a point that is 1 unit away from the origin, the terminal point of an angle is always on the unit circle. (In order to prevent diagrams from becoming too cluttered, we'll usually omit the $(1, 0)$ that indicates that the circle has radius 1.)



Once we have located the terminal point of θ, we define $\cos\theta$, $\sin\theta$, and $\tan\theta$ as follows:

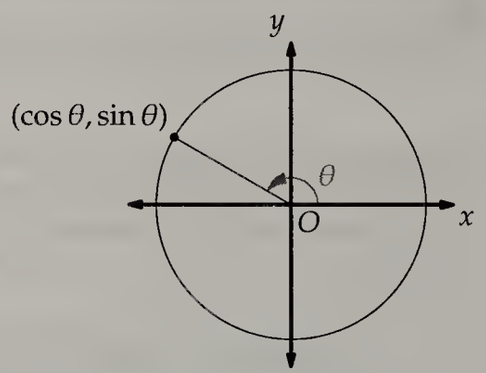
Definition: The **cosine** of an angle θ is the x-coordinate of the terminal point of θ. We denote the cosine of θ as $\cos\theta$.

The **sine** of an angle θ is the y-coordinate of the terminal point of θ. We denote the sine of θ as $\sin\theta$.

The **tangent** of an angle θ is the ratio of the y-coordinate to the x-coordinate of the terminal point of θ. We denote the tangent of θ as $\tan\theta$.

Putting these together, if (x, y) is the terminal point of θ, then

$$\cos\theta = x,$$
$$\sin\theta = y,$$
$$\tan\theta = \frac{y}{x} = \frac{\sin\theta}{\cos\theta}.$$



As the diagram indicates, these definitions can be used for any angle θ, not just acute angles.

One way to remember which coordinate is sine and which is cosine is to remember that it's simply alphabetical order: cosine comes before sine, and x comes before y. To help remember that $\tan\theta$ is the ratio $\frac{y}{x}$ rather than $\frac{x}{y}$, we can think of the tangent of θ as the slope of the line through the origin and the terminal point of θ.

Three other trigonometric functions are **secant**, **cosecant**, and **cotangent**, which are denoted as sec, csc, and cot, respectively. These are defined in terms of cosine, sine, and tangent as follows:

$$\sec\theta = \frac{1}{\cos\theta},$$
$$\csc\theta = \frac{1}{\sin\theta},$$
$$\cot\theta = \frac{1}{\tan\theta} = \frac{\cos\theta}{\sin\theta}.$$

In this section, we become familar with the trigonometric functions by evaluating them at various angles. Feel free to look back at the definitions as you work through the problems. You don't need to memorize them right now; you'll eventually memorize them as you use them.

Finally, as you've probably already realized, "trigonometric" is a long word. Worse yet, it's a mouthful when spoken, so "trigonometry" and "trigonometric" are often shortened as "**trig**."

Problems

Problem 2.1:

(a) Triangle ABC below is a **45-45-90 triangle**, which means one angle is a right angle and the other two angles are equal. If $BC = 1$, then what are AB and AC?

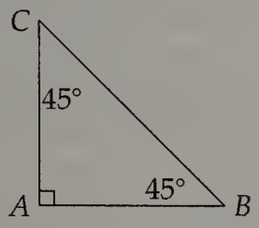


Figure 2.1: Diagram for Part (a)

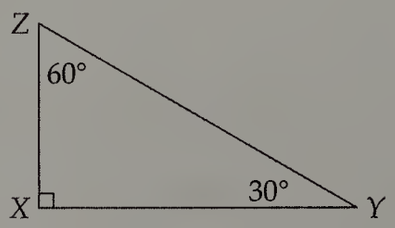


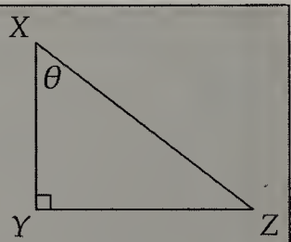
Figure 2.2: Diagram for Part (b)

(b) Triangle XYZ above is a **30-60-90 triangle**, which means that its angles are 30°, 60°, and 90°. If $YZ = 1$, then what are XY and XZ?

Problem 2.2: Find the cosine, sine, and tangent of each of the following angles:

(a)	0°	(c)	45°	(e)	60°
(b)	90°	(d)	30°	(f)	270°

Problem 2.3: The trigonometric functions are often introduced as ratios of sides in a right triangle. For example, in right triangle $\triangle XYZ$ at right, we have $\sin\theta = \frac{YZ}{XZ}$, $\cos\theta = \frac{YX}{XZ}$, and $\tan\theta = \frac{YZ}{XY}$.

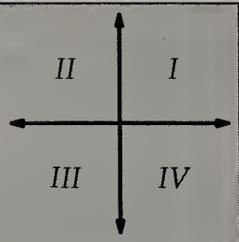


(a) Show that this definition is consistent with the unit circle definition. That is, show that if we compute $\sin\theta$, $\cos\theta$, and $\tan\theta$ using the unit circle, where θ is $\angle X$ in right triangle $\triangle XYZ$ at right, then we get

$$\sin\theta = \frac{YZ}{XZ}, \qquad \cos\theta = \frac{YX}{XZ}, \qquad \tan\theta = \frac{YZ}{YX}.$$

(b) Why are the unit circle definitions more useful than the right triangle definitions?

Problem 2.4: We can think of the coordinate axes as dividing the Cartesian plane into four **quadrants**, which are commonly labeled as shown at right. We say that an angle belongs to the quadrant in which its terminal point lies. So, for example, 120° is in quadrant II and 315° is in quadrant IV. We also say that 120° is a "second quadrant" angle and 315° is a "fourth quadrant" angle. (We typically don't speak of quadrants when referring to an angle whose terminal point is on one of the coordinate axes.)



(a) In which quadrants is sine positive?

(b) In which quadrants is tangent positive?

Problem 2.5: Find each of the following:

(a)	$\cos 120°$	(c)	$\sec 210°$	(e)	$\csc 240°$
(b)	$\sin 135°$	(d)	$\cot 315°$	(f)	$\tan 330°$

Problem 2.6: We can use the unit circle to evaluate trig functions of angles that are more than 360° or less than 0°. If an angle is less than 0°, then we find its terminal point by going *clockwise* around the unit circle from (1, 0), rather than counterclockwise. Find each of the following:

(a) $\cos 1080°$ (c) $\cot(-45°)$

(b) $\sin 3720°$ (d) $\sec(-480°)$

Problem 2.7: We typically write the squares of $\sin x$ and $\cos x$ as $\sin^2 x$ and $\cos^2 x$, and do similarly for other positive powers of trig functions. Prove that for any angle θ, we have $\sin^2\theta + \cos^2\theta = 1$.

Problem 2.8: Find the range of each of the following functions:

(a) $f(t) = \sin t$ (b) $g(t) = \tan t$ (c) $h(t) = \sec t$

Problem 2.1:

(a) Triangle ABC below is a **45-45-90 triangle**, which means one angle is a right angle and the other two angles are equal. If $BC = 1$, then what are AB and AC?

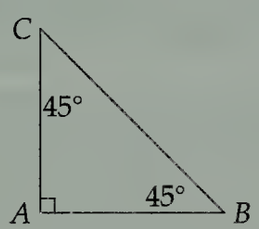


Figure 2.3: Diagram for Part (a)

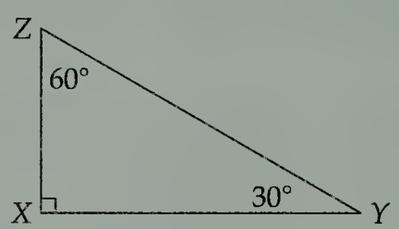


Figure 2.4: Diagram for Part (b)

(b) Triangle XYZ above is a **30-60-90 triangle**, which means that its angles are 30°, 60°, and 90°. If $YZ = 1$, then what are XY and XZ?

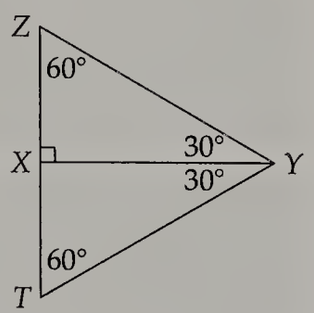
Solution for Problem 2.1:

(a) We let $AB = x$. Since $\angle B = \angle C$, we have $AC = AB = x$. Since $BC = 1$, the Pythagorean Theorem gives us $x^2 + x^2 = 1$, so $x^2 = 1/2$. Therefore, we have

$$x = \sqrt{\frac{1}{2}} = \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}} \cdot \frac{\sqrt{2}}{\sqrt{2}} = \frac{\sqrt{2}}{2}.$$

Both $\overline{AB}$ and $\overline{AC}$ have length $\sqrt{2}/2$.

(b) The 60° and 30° angles makes us think of equilateral triangles. By reflecting $\triangle XYZ$ over side $\overline{XY}$, we form equilateral triangle TYZ in the diagram at right. Since $\triangle TYZ$ is equilateral, we have $TZ = ZY = 1$. Since $\triangle TYX \cong \triangle ZYX$, we have $TX = XZ$. Putting this together with $TZ = 1$ gives $XZ = 1/2$. Finally, the Pythagorean Theorem applied to $\triangle XYZ$ gives



$$XY = \sqrt{YZ^2 - XZ^2} = \sqrt{1 - \frac{1}{4}} = \sqrt{\frac{3}{4}} = \frac{\sqrt{3}}{2}.$$

□

Important: If the hypotenuse of an isosceles right triangle has length 1, then each leg has length $\frac{\sqrt{2}}{2}$. If the acute angles of a right triangle are 30° and 60° and the hypotenuse has length 1, then the leg opposite the 30° angle has length $\frac{1}{2}$ and the leg opposite the 60° angle has length $\frac{\sqrt{3}}{2}$. These relationships are depicted below.

Problem 2.2: Find the cosine, sine, and tangent of each of the following angles:

(a) 0° (c) 45° (e) 60°

(b) 90° (d) 30° (f) 270°

Solution for Problem 2.2:

(a) The point on the unit circle that is 0° counterclockwise from (1,0) is simply (1,0). Since the terminal point of 0° is (1,0), we have

$$\cos 0° = 1,$$
$$\sin 0° = 0,$$
$$\tan 0° = 0.$$

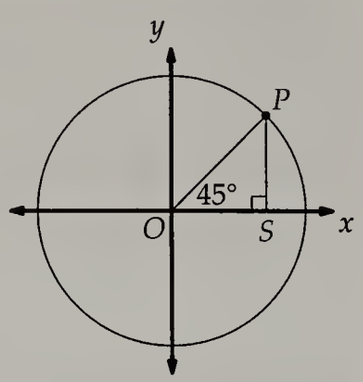
(b) The point on the unit circle that is 90° counterclockwise from (1,0) is (0,1). Since the terminal point of 90° is (0,1), we have $\cos 90° = 0$ and $\sin 90° = 1$.

Since $\tan\theta$ is defined as $\dfrac{\sin\theta}{\cos\theta}$, $\tan\theta$ is undefined whenever $\cos\theta = 0$.

Important: If $\cos\theta = 0$, then $\tan\theta$ is undefined.

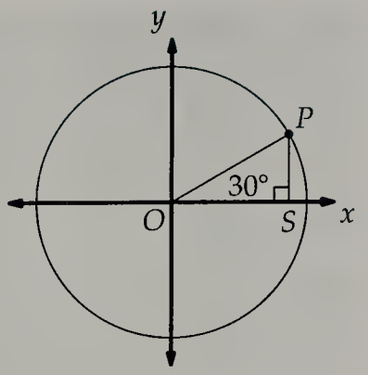
Therefore, $\tan 90°$ is undefined, because $\cos 90° = 0$. We can also see this in terms of the unit circle. The slope of the line from the origin to (0,1) is undefined, so $\tan 90°$ is undefined. Similarly, $\tan\theta$ is undefined whenever the terminal point of θ is on the y-axis (and therefore has x-coordinate 0, which means that $\cos\theta$ is 0).

(c) We first find the terminal point, P, of 45° by rotating the point (1,0) counterclockwise an angle of 45°. To find the coordinates of P, we build right triangle OPS, as shown, and see that $\triangle OPS$ is a 45-45-90 right triangle with hypotenuse $OP = 1$. Therefore, we have $OS = SP = \frac{\sqrt{2}}{2}$, so point P is $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ and:



$$\cos 45° = \frac{\sqrt{2}}{2},$$
$$\sin 45° = \frac{\sqrt{2}}{2},$$
$$\tan 45° = 1.$$

(d) As in the previous part, we locate the terminal point, P, and then build a right triangle by drawing the altitude from P to the x-axis. Since $\triangle OPS$ is a 30-60-90 triangle with $OP = 1$ and $\angle POS = 30°$, we have $OS = \frac{\sqrt{3}}{2}$ and $PS = \frac{1}{2}$, so point P is $\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$. Therefore, we have

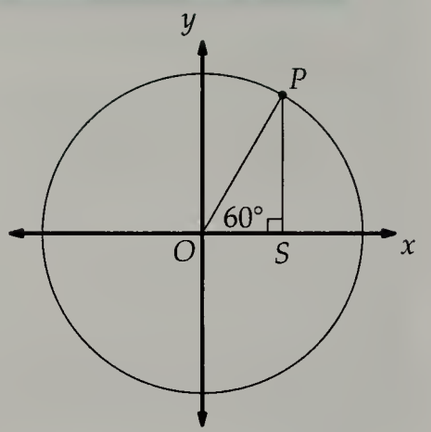


$$\cos 30° = \frac{\sqrt{3}}{2},$$
$$\sin 30° = \frac{1}{2},$$
$$\tan 30° = \frac{1/2}{\sqrt{3}/2} = \frac{1}{\sqrt{3}} = \frac{1}{\sqrt{3}} \cdot \frac{\sqrt{3}}{\sqrt{3}} = \frac{\sqrt{3}}{3}.$$

The last two parts show a common strategy for evaluating trigonometric functions of some angles:

Important: To evaluate trigonometric functions of angles that are multiples of 30° or 45°, we can locate the terminal point of the angle on the unit circle, and then build a right triangle by drawing an altitude from the terminal point to the x-axis.

(e) Once again, we build a right triangle after locating the terminal point. Since $OP = 1$ and $\angle POS = 60°$, we have $OS = \frac{1}{2}$ and $PS = \frac{\sqrt{3}}{2}$. Therefore, point P is $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$, and we have



$$\cos 60° = \frac{1}{2},$$
$$\sin 60° = \frac{\sqrt{3}}{2},$$
$$\tan 60° = \frac{\sqrt{3}/2}{1/2} = \sqrt{3}.$$

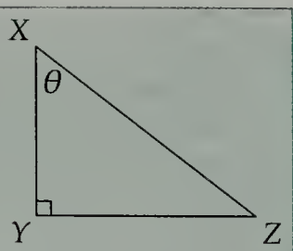
(f) The point $(0, -1)$ is 270° counterclockwise from $(1, 0)$. Therefore, we have

$$\cos 270° = 0,$$
$$\sin 270° = -1,$$
$$\tan 270° \text{ is undefined.}$$

□

This book may not be the first time you have encountered trigonometric functions. Trig functions are often introduced in geometry classes, where they are defined a little differently.

Problem 2.3: The trigonometric functions are often introduced as ratios of sides in a right triangle. For example, in right triangle $\triangle XYZ$ at right, we have $\sin\theta = \frac{YZ}{XZ}$, $\cos\theta = \frac{YX}{XZ}$, and $\tan\theta = \frac{YZ}{XY}$.



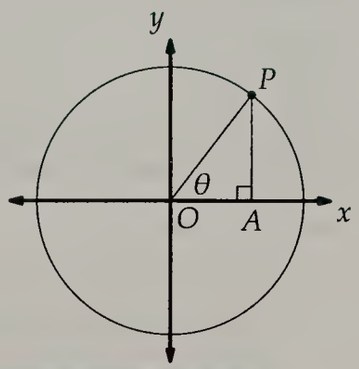
(a) Show that this definition is consistent with the unit circle definition. That is, show that if we compute $\sin\theta$, $\cos\theta$, and $\tan\theta$ using the unit circle, where θ is $\angle X$ in right triangle $\triangle XYZ$ above, then we get

$$\sin\theta = \frac{YZ}{XZ}, \qquad \cos\theta = \frac{YX}{XZ}, \qquad \tan\theta = \frac{YZ}{YX}.$$

(b) Why are the unit circle definitions more useful than the right triangle definitions?

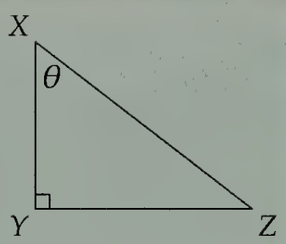
Solution for Problem 2.3:

(a) We start by finding the terminal point of θ on the unit circle. From the definition of $\sin\theta$, we have $\sin\theta = PA$. To relate this to $\triangle XYZ$ in the problem, we note that $\angle X = \angle POA$ and $\angle Y = \angle PAO$, so $\triangle XYZ \sim \triangle OAP$, which means that $\frac{YZ}{XZ} = \frac{PA}{PO}$. Since $PO = 1$, we have $\frac{YZ}{XZ} = PA$. Therefore, the unit circle definition gives the same result for $\sin\theta$ as the right triangle definition.



Similarly, we have $\cos\theta = OA$ and $\tan\theta = \frac{PA}{OA}$. From $\triangle XYZ \sim \triangle OAP$, we have $\frac{YX}{XZ} = \frac{OA}{PO} = OA$ and $\frac{YZ}{YX} = \frac{PA}{OA}$, so the right triangle definitions of $\cos\theta$ and $\tan\theta$ for acute angles θ are consistent with their unit circle definitions.

Important: If triangle XYZ is a right triangle with $\angle XYZ = 90°$ and $\angle X = \theta$, then



$$\sin\theta = \frac{YZ}{XZ}, \qquad \cos\theta = \frac{YX}{XZ}, \qquad \tan\theta = \frac{YZ}{YX}.$$

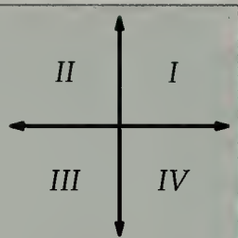
Another way of viewing this is: if θ is the measure of an acute angle in a right triangle, then we have

$$\sin\theta = \frac{\text{opposite leg}}{\text{hypotenuse}}, \qquad \cos\theta = \frac{\text{adjacent leg}}{\text{hypotenuse}}, \qquad \tan\theta = \frac{\text{opposite leg}}{\text{adjacent leg}}.$$

(b) The unit circle definition is much more powerful because it allows us to define trig functions for non-acute angles.

□

Problem 2.4: We can think of the coordinate axes as dividing the Cartesian plane into four **quadrants**, which are commonly labeled as shown at right. We say that an angle belongs to the quadrant in which its terminal point lies. So, for example, 120° is in quadrant II and 315° is in quadrant IV. We also say that 120° is a "second quadrant" angle and 315° is a "fourth quadrant" angle. (We typically don't speak of quadrants when referring to an angle whose terminal point is on one of the coordinate axes.)



(a) In which quadrants is sine positive?

(b) In which quadrants is tangent positive?

Solution for Problem 2.4:

(a) The sine of an angle is positive if and only if the y-coordinate of the angle's terminal point is positive. Therefore, sine is positive in quadrants I and II, and negative in quadrants III and IV.

(b) Since tangent is the quotient of sine and cosine, it is positive if and only if sine and cosine are both positive or both negative. Sine and cosine are both positive in the first quadrant, and both negative in the third quadrant (since both x and y are negative in the third quadrant). Therefore, tangent is positive in quadrants I and III. In quadrant II, the x-coordinate is negative while the y-coordinate is positive, so tangent is negative. In quadrant IV, the x-coordinate is positive while the y-coordinate is negative, so tangent is again negative.

□

You don't have to memorize which trig functions are positive and which are negative in each quadrant. This information will come naturally if you understand the unit circle and the definitions of the trig functions. As practice, determine the sign of the answers in the following problems *before finding the answers*.

Problem 2.5: Find each of the following:

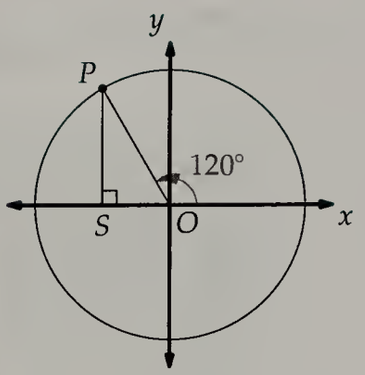
(a) $\cos 120°$ (d) $\cot 315°$

(b) $\sin 135°$ (e) $\csc 240°$

(c) $\sec 210°$ (f) $\tan 330°$

Solution for Problem 2.5:

(a) We first find the terminal point of 120°. Since 120° is between 90° and 180°, the terminal point of 120° is in the second quadrant, as indicated by point P. P is $120° - 90° = 30°$ beyond the y-axis, so when we draw altitude $\overline{PS}$, we have $\angle POS = 90° - 30° = 60°$. Therefore, $\triangle POS$ is a 30-60-90 triangle with $OS = \frac{1}{2}$ and $PS = \frac{\sqrt{3}}{2}$. We then must be careful not make the following mistake:



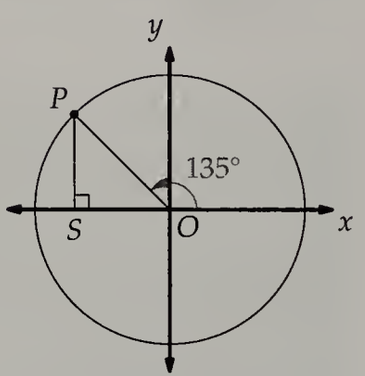
Bogus Solution: The coordinates of P are $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$, so

$$\cos 120° = \frac{1}{2}.$$

The mistake here is that P is to the left of the y-axis. Therefore, while $OS = \frac{1}{2}$, the x-coordinate of P is $-\frac{1}{2}$, not $\frac{1}{2}$. So, we have

$$\cos 120° = -\frac{1}{2}.$$

(b) Since 135° is between 90° and 180°, it is a second quadrant angle. The terminal point, P, is $180° - 135° = 45°$ clockwise from $(-1, 0)$, so drawing altitude $\overline{PS}$ forms a 45-45-90 right triangle, from which we find $PS = \frac{\sqrt{2}}{2}$. Since P is above the x-axis, its y-coordinate is positive, and we have

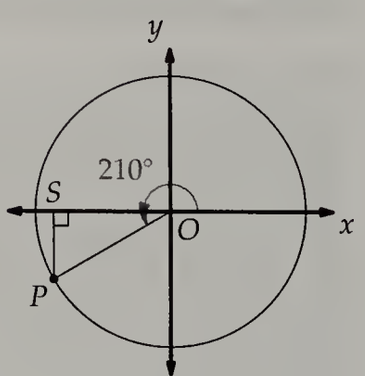


$$\sin 135° = \frac{\sqrt{2}}{2}.$$

(c) First, we note that

$$\sec 210° = \frac{1}{\cos 210°},$$

so we must find the x-coordinate of the terminal point, P, of 210°. P is $210° - 180° = 30°$ counterclockwise from $(-1, 0)$, so it is in the third quadrant. Drawing altitude $\overline{PS}$ to the x-axis forms 30-60-90 triangle POS with $\angle POS = 30°$, so $OS = \frac{\sqrt{3}}{2}$. Since P is to the left of the y-axis, the x-coordinate of P is $-OS$, which is $-\frac{\sqrt{3}}{2}$. So, we have

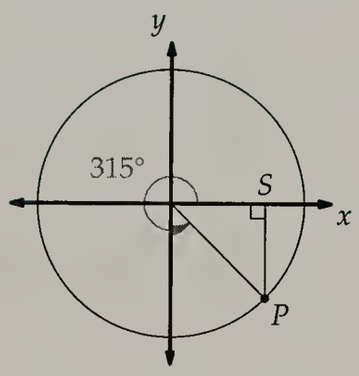


$$\sec 210° = \frac{1}{\cos 210°} = \frac{1}{-\sqrt{3}/2} = -\frac{2}{\sqrt{3}} = -\frac{2\sqrt{3}}{3}.$$

(d) First, we note that

$$\cot 315^\circ = \frac{1}{\tan 315^\circ}.$$

Since 315° is 45° less than 360°, its terminal point, P, is just 45° less than a full counterclockwise revolution of $(1,0)$ about the origin. Therefore, P is in the fourth quadrant, as shown, and $\angle POS = 45^\circ$. This gives us $OS = SP = \frac{\sqrt{2}}{2}$. Because P is to the right of the y-axis and below the x-axis, its coordinates are $\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$. So, we have



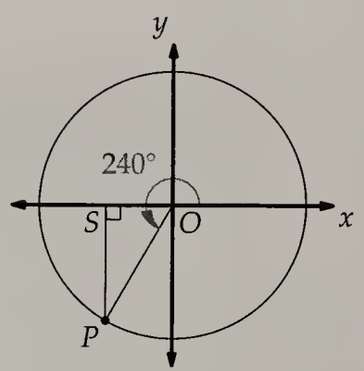
$$\tan 315^\circ = \frac{-\sqrt{2}/2}{\sqrt{2}/2} = -1,$$

which gives us $\cot 315^\circ = 1/(\tan 315^\circ) = -1$.

(e) We have

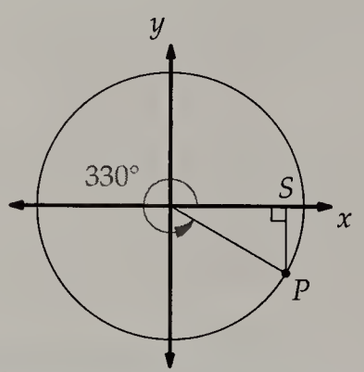
$$\csc 240^\circ = \frac{1}{\sin 240^\circ},$$

so we seek the y-coordinate of the terminal point of 240°. Since 240° is between 180° and 270°, it is a third quadrant angle whose terminal point is $240^\circ - 180^\circ = 60^\circ$ counterclockwise from $(-1,0)$. In right triangle POS, we have $\angle POS = 60^\circ$, so $OS = \frac{1}{2}$ and $PS = \frac{\sqrt{3}}{2}$. Because P is below the x-axis, its y-coordinate is negative:



$$\csc 240^\circ = \frac{1}{\sin 240^\circ} = \frac{1}{-\sqrt{3}/2} = -\frac{2}{\sqrt{3}} = -\frac{2\sqrt{3}}{3}.$$

(f) Since 330° is 30° less than 360°, the terminal point of 330° is 30° *clockwise* from $(1,0)$, which puts it in the fourth quadrant with $\angle POS = 30^\circ$, as shown. We therefore have $OS = \frac{\sqrt{3}}{2}$ and $PS = \frac{1}{2}$, so P is $\left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right)$, and



$$\tan 330^\circ = \frac{-1/2}{\sqrt{3}/2} = -\frac{1}{\sqrt{3}} = -\frac{\sqrt{3}}{3}.$$

□

We can also use the unit circle to evaluate trig functions of angles that are more than 360° or less than 0°. If an angle is less than 0°, then we find its terminal point by going *clockwise* around the unit circle from $(1,0)$, rather than counterclockwise.

Problem 2.6: Find each of the following:

(a) $\cos 1080^\circ$

(b) $\sin 3720^\circ$

(c) $\cot(-45^\circ)$

(d) $\sec(-480^\circ)$

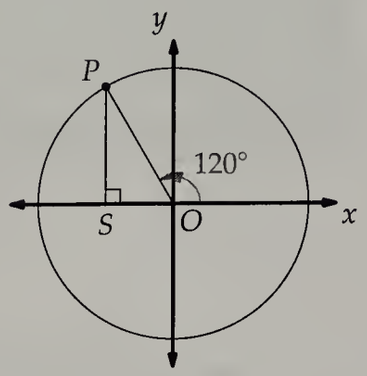
Solution for Problem 2.6:

(a) Going 1080° counterclockwise around the unit circle from $(1,0)$ gets us to the same terminal point as going $1080^\circ - 360^\circ = 720^\circ$ counterclockwise, since 360° is simply one full revolution around the circle. Since $720^\circ = 2 \cdot 360^\circ$, going 720° around the unit circle is two full revolutions, which brings us back to $(1,0)$ as the terminal point. Therefore, we have:

$$\cos 1080^\circ = \cos 720^\circ = \cos 0^\circ = 1.$$

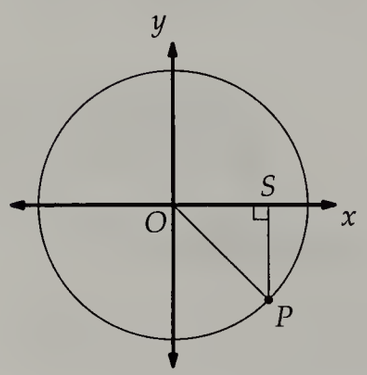
We can similarly evaluate trig functions of any angle greater than 360° by subtracting multiples of 360° until we have an angle that is less than 360°.

(b) Rather than repeatedly subtracting 360°, we notice that 3720° is slightly greater than $10 \cdot 360°$, so we subtract 3600°, which is the same as 10 full revolutions around the unit circle. Therefore, 3720° has the same terminal point as the angle $3720° - 10 \cdot 360° = 120°$. So, we have $\sin 3720° = \sin 120°$. In the diagram at right, we find the terminal point, P, of 120°, and form the 30-60-90 triangle POS with $\angle POS = 60°$. Since $PS = \frac{\sqrt{3}}{2}$, and P is above the x-axis, we have



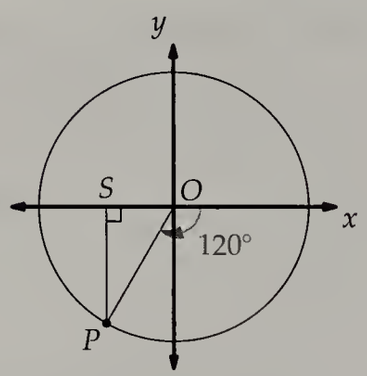
$$\sin 3720° = \sin 120° = \frac{\sqrt{3}}{2}.$$

(c) Since the angle is negative, the terminal point is 45° *clockwise* from $(1, 0)$, as shown. From 45-45-90 triangle POS, we have $PS = OS = \frac{\sqrt{2}}{2}$, so the coordinates of P are $\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$, which gives us



$$\cot(-45°) = \frac{1}{\tan(-45°)} = \frac{\cos(-45°)}{\sin(-45°)} = \frac{\sqrt{2}/2}{-\sqrt{2}/2} = -1.$$

(d) Going 480° clockwise around the unit circle brings us to the same terminal point as going $480° - 360° = 120°$ clockwise. This is the same as noting that the terminal point of −480° is the same as the terminal point of $-480° + 360° = -120°$. We find the terminal point by going 120° *clockwise* about the unit circle from $(1, 0)$, or by noting that the terminal point of −120° is the same as the terminal point of $-120° + 360° = 240°$. Either way, we have 30-60-90 triangle POS with $\angle POS = 60°$, from which we find $OS = \frac{1}{2}$. Since P is to the left of the y-axis, its x-coordinate is negative, and we have



$$\sec(-480°) = \sec(-120°) = \frac{1}{\cos(-120°)} = \frac{1}{-1/2} = -2.$$

□

Understanding the unit circle not only allows us to evaluate trigonometric functions of some angles, but it also allows us to visualize and sometimes prove **trigonometric identities**. A trigonometric identity is an equation that is true for all angles for which the equation is defined. We now introduce the most common trig identity; we'll explore trig identities in much greater detail in Chapter 3.

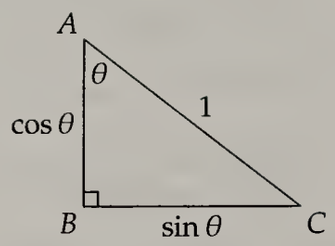
Problem 2.7: We typically write the squares of $\sin x$ and $\cos x$ as $\sin^2 x$ and $\cos^2 x$, and do similarly for other positive powers of trig functions. Prove that for any angle θ, we have $\sin^2 \theta + \cos^2 \theta = 1$.

Solution for Problem 2.7: By definition, the point $(x, y) = (\cos\theta, \sin\theta)$ is the point on the unit circle that is θ degrees counterclockwise about the origin from $(1, 0)$. The unit circle is the graph of the equation $x^2 + y^2 = 1$. Since the point $(\cos\theta, \sin\theta)$ is on this graph, the ordered pair $(x, y) = (\cos\theta, \sin\theta)$ must satisfy the equation $x^2 + y^2 = 1$, which gives us $\cos^2\theta + \sin^2\theta = 1$. Reversing the addends on the left gives us the more commonly written form of this trigonometric identity:

Important: For any θ, we have $\sin^2\theta + \cos^2\theta = 1$.

□

The identity $\sin^2\theta + \cos^2\theta = 1$ isn't really something new—it's essentially the Pythagorean Theorem. At right is a right triangle ABC with hypotenuse of length 1 and $\angle BAC = \theta$. Because $\sin\theta = \frac{BC}{AC}$ and $AC = 1$, we have $BC = \sin\theta$. Similarly, $\cos\theta = \frac{AB}{AC}$, so $AB = \cos\theta$. The Pythagorean Theorem gives us



$$BC^2 + AB^2 = AC^2,$$

so

$$\sin^2\theta + \cos^2\theta = 1.$$

WARNING!! In Section 1.3, we noted that $f^2(x)$ usually means $f(f(x))$. Trigonometric functions are a common exception to this. When we write $\sin^2 x$, we mean $(\sin x)^2$, not $\sin(\sin x)$.

Problem 2.8: What is the range of each of the following functions:

(a) $f(t) = \sin t$

(b) $g(t) = \tan t$

(c) $h(t) = \sec t$

Solution for Problem 2.8:

(a) Any point (x, y) on the Cartesian plane with a y-coordinate that is greater than 1 or less than -1 must be more than 1 unit from the origin (since $\sqrt{x^2 + y^2} > 1$ if $|y| > 1$ and x is real). Therefore, no point on the unit circle can have y-coordinate greater than 1 or less than -1, which means $-1 \le \sin t \le 1$. Every value in the interval $[-1, 1]$ is in the range of $\sin t$, since the line $y = k$ intersects the unit circle for any value of k such that $-1 \le k \le 1$. So, the range of $f(t) = \sin t$ is $[-1, 1]$. Similarly, the range of $\cos t$ is also $[-1, 1]$.

Important: For any angle θ, we have $-1 \le \sin\theta \le 1$ and $-1 \le \cos\theta \le 1$.

(b) The value of $\tan t$ is the slope of the line between the origin and the terminal point of the angle t. Therefore, for any real number m, we can find a value of t such that $\tan t = m$ by constructing the line through the origin with slope m. This line must intersect the unit circle, and each point at which the line intersects the unit circle is the terminal point of an angle whose tangent is m. Therefore, the range of $g(t) = \tan t$ is all real numbers.

(c) First, we note that $\sec t = 1/(\cos t)$, so we can find the range of $\sec t$ from the range of $\cos t$, which is $[-1, 1]$. If we have $0 < k \le 1$, then the reciprocal of k is at least 1. Conversely, any number greater than or equal to 1 is the reciprocal of a number in the range of $\cos t$. Therefore, every number in the interval $[1, +\infty)$ is in the range of $h(t) = \sec t$. Similarly, every number in the interval $(-\infty, -1]$ is in the range of h. Finally, no number in the interval $(-1, 1)$ is in the range of $\sec t$, since such a number being in the range of $\sec t$ would mean that $|\cos t| > 1$, which is impossible. Combining these observations, the range of $\sec t$ is $(-\infty, -1] \cup [1, +\infty)$.

□

Exercises

2.1.1 Evaluate each of the following:

(a) $\sin 120°$ (c) $\sec 45°$ (e) $\cos 180°$

(b) $\cot 60°$ (d) $\tan 210°$ (f) $\csc 300°$

2.1.2 Evaluate each of the following:

(a) $\cos(-90°)$ (c) $\sin 750°$

(b) $\tan 36000°$ (d) $\sec(-1200°)$

2.1.3 Which of the following are positive: $\sin 293°$, $\cos(-68°)$, $\cot 206°$, $\csc 90.5°$?

2.1.4 For how many values of θ such that $0 \le \theta \le 900°$ is $\cos\theta = \frac{\sqrt{2}}{2}$?

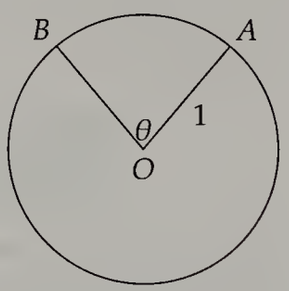
2.1.5 What is the range of the function $f(x) = 7 + 2\sin 3x$?

2.2 Radians

There are many different units we can use to express distance. Our choice of units is usually determined by what we are measuring. For example, we wouldn't measure the distance between two cities in centimeters or inches; we usually use kilometers or miles. Similarly, we use feet or meters for the height of a person instead of miles or kilometers.

Just as with distance, there are different units for angles, and our choice of units is often determined by our purpose. In geometry, we typically use degrees, which we denote with the usual degree symbol, °. However, trigonometric functions of angles are a significant part of algebra, and in an algebraic context, degrees are not a convenient unit. The most convenient unit of measure for angles is the **radian**.

To measure an angle in radians, we draw a circle with radius 1 centered at the vertex of the angle. The measure of the angle in radians equals the length of the arc cut off by the angle. So, in the diagram at right the measure of θ in radians equals the length of arc $\overset{\frown}{AB}$. As another example, the measure of a right angle in radians is $\frac{\pi}{2}$ because the circumference of a circle with radius 1 is 2π and a right angle cuts off one-quarter of the circumference.



While we introduced the idea of a different unit for angles by mentioning the many units available for length, there is a much more fundamental difference between degrees and radians than there is between inches and centimeters. The choice of what distance is "one inch" is an arbitrary human creation. Similarly, the choice of what angle measure is "one degree" is arbitrary because we could have chosen any number of degrees to be the number of degrees in a circle.

With radians, the situation is different. By specifying the measure of an angle as the arc length cut off by the angle in a circle of radius 1, we are really defining a radian with a *ratio*. We could have equivalently defined the measure of an angle in radians as "the ratio of the length of the arc cut off by the angle to the radius." This means that the radian is a **dimensionless unit**. By this, we mean that we do not use some human-defined angle as a reference to tell us what a radian is. For degrees, we need "a full circle is 360°" to tell us what a degree is, so the degree is not a dimensionless unit. For meters, we use the reference length "the distance traveled by light in a vacuum during 1/299792458 of a second," so the meter is not a dimensionless unit. (No, we didn't make that number up!) But to measure an angle in radians, we don't need to compare the angle to any human-defined reference point, so the radian is a dimensionless unit.

Because the radian is a dimensionless unit, we usually use it instead of degrees when working with an angle measure in an algebraic context. For example, suppose the measure of an angle is x, and we must work in a problem with the quantity x^2. Measuring the angle in degrees makes x^2 puzzling: what is a square degree? Measuring the angle in radians makes more sense; the measure is a number without any dimension, so squaring it is fine—that just gives us another number without any dimension. As we'll see later in the text, there are times that we wish to raise a number to a power that involves an angle measure. Raising a number to dimensionless number makes sense, but what is $2^{3\text{ inches}}$? What units would such a result have? There's no good answer to this question, so we generally don't like to raise numbers to powers that are expressed in terms of a unit that is not dimensionless.

Throughout this text, if we are evaluating a trigonometric function of a number that is without units, then the number should be assumed to be in radians, not in degrees. Therefore, $\sin 1$ does *not* mean the same thing as $\sin 1°$. The expression $\sin 1$ is the sine of 1 radian, not 1 degree.

Problems

Problem 2.9: Convert each of the following degree measures to radians:

(a) $60°$ (c) $-315°$

(b) $288°$ (d) $1400°$

Problem 2.10: Convert each of the following measures in radians to degrees:

(a) $2\pi/5$ (c) $-7\pi/3$

(b) $4\pi/3$ (d) 1

Problem 2.11: Evaluate each of the following:

(a) $\cos \frac{2\pi}{3}$ (c) $\tan(-11\pi)$

(b) $\sin \frac{7\pi}{3}$ (d) $\sec \frac{3\pi}{4}$

Problem 2.12: What is the sign of $\cot 30$?

Problem 2.9: Convert each of the following degree measures to radians:

(a) $60°$ (c) $-315°$

(b) $288°$ (d) $1400°$

Solution for Problem 2.9:

(a) Since $60°$ is $60°/360° = 1/6$ of a full circle, its equivalent in radians is $(1/6)(2\pi) = \pi/3$ radians.

(b) Since $288°$ is $288°/360° = 4/5$ of a full circle, its equivalent in radians is $(4/5)(2\pi) = 8\pi/5$ radians.

(c) Since $315°$ is $315°/360° = 7/8$ of a full circle, the equivalent of $-315°$ in radians is $(7/8)(-2\pi) = -7\pi/4$ radians.

(d) Since $1400°$ is $1400°/360° = 35/9$ of a full circle, its equivalent in radians is $(35/9)(2\pi) = 70\pi/9$ radians.

□

Rather than using the two-step process we used in our solutions above, we can use a **conversion factor** to do

both steps at once. Specifically, since 360° is equivalent to 2π radians, we can use the shorthand

$$\frac{2\pi}{360^\circ} = 1.$$

As an example, to convert 288° to radians, we multiply 288° by the conversion factor. This leaves the value unchanged because the conversion factor equals 1:

$$288^\circ \cdot \frac{2\pi}{360^\circ} = \frac{288^\circ}{360^\circ} \cdot 2\pi = \frac{4}{5} \cdot 2\pi = \frac{8\pi}{5}.$$

Notice how the degrees "cancel out."

Converting radians to degrees is often even easier. Because 2π radians equals 360°, we know that π radians is equivalent to 180°. So, in an expression in radians that includes π, we can simply replace π with 180°.

Problem 2.10: Convert each of the following measures in radians to degrees:

(a) $2\pi/5$ (c) $-7\pi/3$

(b) $4\pi/3$ (d) 1

Solution for Problem 2.10:

(a) Since π radians is 180°, we know that $2\pi/5$ radians equals $2(180^\circ)/5$, which equals 72°.

(b) Since π radians is 180°, we know that $4\pi/3$ radians equals $4(180^\circ)/3$, which equals 240°.

(c) As before, we replace π radians with 180°, so $-7\pi/3$ radians equals $-7(180^\circ)/3$, which equals −420°.

(d) Oh no! We don't have π in this expression. We don't have to use π when expressing angles in terms of radians. Fortunately, we can still use our conversion factor approach by simply inverting our earlier conversion factor:

$$\frac{360^\circ}{2\pi} = 1.$$

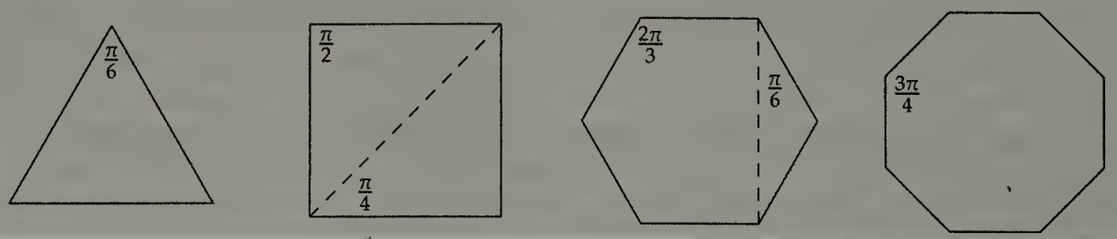
Now, when we multiply our angle in radians by this conversion factor, we are left with a measure in degrees:

$$1 \cdot \frac{360^\circ}{2\pi} = \frac{180^\circ}{\pi}.$$

Notice that we don't sprinkle the word "radians" throughout this computation. Typically, we don't write the units with angle measures expressed in radians, and any angle measure written without units is assumed to be in radians, not degrees.

□

Extra! Some common angle measures in radians are indicated in the regular polygons shown below.



Problem 2.11: Evaluate each of the following:

(a) $\cos\frac{2\pi}{3}$ (c) $\tan(-11\pi)$

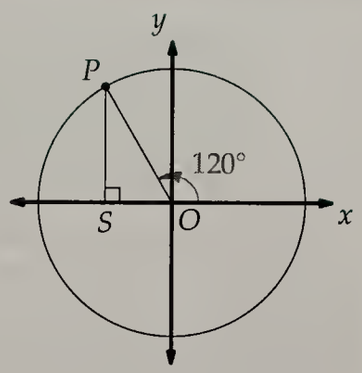
(b) $\sin\frac{7\pi}{3}$ (d) $\sec\frac{3\pi}{4}$

Solution for Problem 2.11:

(a) We first convert radians to degrees. Since π radians is equivalent to 180°, we have

$$\cos\frac{2\pi}{3} = \cos\frac{2(180°)}{3} = \cos 120°.$$

Now, we can proceed as usual, by locating the terminal point of 120° on the unit circle as shown at right. From 30-60-90 triangle POS, we find $OS = 1/2$. Since P is to the left of the y-axis, we have $\cos\frac{2\pi}{3} = \cos 120° = -\frac{1}{2}$.



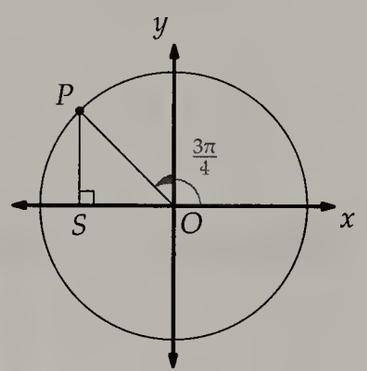
(b) Since $\frac{7\pi}{3} = \frac{7(180°)}{3} = 420°$, we have

$$\sin\frac{7\pi}{3} = \sin 420° = \sin(420° - 360°) = \sin 60° = \frac{\sqrt{3}}{2}.$$

(c) Rather than converting to degrees first, we instead note that because a full revolution around the unit circle is 2π radians, we have $\tan(-11\pi) = \tan(-11\pi + 5\cdot 2\pi) = \tan(-\pi)$. Since a full revolution is 2π, a half revolution is π, so the terminal point of an angle of $-\pi$ radians is $(-1,0)$. Therefore, $\sin(-\pi) = 0$, which means

$$\tan(-11\pi) = \tan(-\pi) = \frac{\sin(-\pi)}{\cos(-\pi)} = 0.$$

(d) We could convert the angle to degrees, or we could note that the terminal point of $\frac{\pi}{2}$ radians is $(0,1)$ and the terminal point of π radians is $(-1,0)$. Since $\frac{3\pi}{4}$ radians is the average of these two angles, its terminal point is the midpoint of the arc connecting these angles' terminal points. In other words, its terminal point is the second quadrant point P shown in the diagram at right, in which $\angle POS = \frac{\pi}{4} = 45°$. From 45-45-90 $\triangle POS$, we have $OS = PS = \frac{\sqrt{2}}{2}$. What's wrong with this finish:



Bogus Solution: We therefore have

$$\sec\frac{3\pi}{4} = \frac{1}{\sin\frac{3\pi}{4}} = \frac{1}{\sqrt{2}/2} = \frac{2}{\sqrt{2}} = \sqrt{2}.$$

The error here is that $\sec x = \dfrac{1}{\cos x}$, not $\dfrac{1}{\sin x}$.

WARNING!! It's very easy to confuse the definitions of $\sec x$ and $\csc x$. Don't be afraid to look them up when you encounter secant and cosecant.

Correcting this mistake, we have

$$\sec\frac{3\pi}{4} = \frac{1}{\cos\frac{3\pi}{4}} = \frac{1}{-\sqrt{2}/2} = -\frac{2}{\sqrt{2}} = -\sqrt{2}.$$

□

Problem 2.12: What is the sign of $\cot 30$?

Solution for Problem 2.12: What's wrong with this solution:

Bogus Solution: We have $\cot 30 = \frac{\cos 30}{\sin 30}$. Since 30 is between 0 and 90, it is a first quadrant angle, which means that both $\cos 30$ and $\sin 30$ are positive. Therefore, $\cot 30$ is also positive.

The reasoning in this solution would be correct if the question had asked for the sign of $\cot 30°$. However, the problem asks for $\cot 30$, not $\cot 30°$.

WARNING!! If an angle measure does not include the degree symbol or any indication of units, then the measure is assumed to be in *radians*, not degrees. So, if you are asked to find $\sin 30$, the answer is not $1/2$, since $\sin 30$ refers to the sine of an angle of 30 radians, not 30 degrees.

In order to find the sign of $\cot 30$, we must find the quadrant of an angle of 30 radians. We can take a bit of a shortcut by noticing that π is approximately 3.14, so 10π, which has terminal point $(1,0)$, is approximately 31.4 radians. Therefore, an angle of 30 radians is approximately $31.4 - 30 = 1.4$ radians *clockwise* from $(1,0)$. (Remember, 30 radians is slightly *less* than 10π radians, so its terminal point will be less than 5 full revolutions counterclockwise from $(1,0)$.) But how large is an angle of 1.4 radians?

Again, we can estimate. Since π is approximately 3.14, we know that $\frac{\pi}{2}$ radians, which is 90°, is approximately 1.57 radians. Therefore, 1.4 radians is less than 90°. Since 30 radians is approximately 1.4 radians clockwise from $(1,0)$, and 1.4 radians is less than 90°, we know that the terminal point of 30 radians is in the fourth quadrant. This means that $\cos 30$ is positive and $\sin 30$ is negative, so $\cot 30$ is negative. □

WARNING!! When using a calculator or computer, make sure you know whether your device or program is using degrees or radians. If it assumes that the angle is in radians, and you enter "tan(30)" to find $\tan 30°$, the answer will not be what you intend. One quick way to test it is to use it to compute $\sin 30°$. If it returns something like -0.98803 instead of 0.5, then it is expecting the angles in radians, not degrees.

Exercises

2.2.1 Convert each of the following degree measures to radians:

(a) 330° (c) −315°

(b) 270° (d) 2000°

2.2.2 Convert each of the following measures in radians to degrees:

(a) 3π (c) $-7\pi/5$

(b) $2\pi/3$ (d) -3

2.2.3 Draw a 30-60-90 triangle and a 45-45-90 triangle. In each triangle, label the angles with their measures in radians. Let the hypotenuse of each triangle be 1, and label the side lengths appropriately.

2.2.4 Evaluate each of the following:

(a) $\sin \frac{\pi}{4}$ (c) $\tan \frac{5\pi}{6}$

(b) $\csc 7\pi$ (d) $\cos\left(-\frac{\pi}{3}\right)$

2.2.5 What is the domain of the function $f(x) = 3\tan\left(2x - \frac{\pi}{3}\right)$?

2.2.6 Jake was taking a trigonometry test, and was asked to find $\tan 40°$. He used his calculator to find the answer by typing in 40 and then hitting the tangent key. He dutifully wrote down what his calculator told him, -1.11721. He was very upset to have his answer marked incorrect, and was convinced that his calculator was broken.

(a) Even without knowing what $\tan 40°$ equals, how should Jake have immediately known that his answer of -1.11721 was not the correct answer to the problem?

(b) What mistake did he probably make?

2.3 Graphs of Trigonometric Functions

In this section, we study the graphs of trigonometric functions. Throughout our discussion of graphing trig functions, all angles will be expressed in radians.

Problems

Problem 2.13:

(a) Find the graph of $y = \sin x$ for $0 \leq x \leq 2\pi$.

(b) Use your answer to part (a) to construct the graph of $y = \sin x$ for all x.

(c) A function f is called **periodic** if there is some nonzero constant t such that $f(x) = f(x + t)$ for all values of x such that $f(x)$ and $f(x + t)$ are defined. The smallest possible positive value of t is called the **period** of f (if such a smallest value exists). Show that $g(x) = \sin x$ is periodic and find its period.

Problem 2.14: Graph $y = \cos x$. What is the period of $\cos x$?

Problem 2.15: What is the relationship between the graph of $y = \sin x$ and the graph of $y = \cos x$? What does this suggest as a relationship between sine and cosine?

Problem 2.16:

(a) The graph of $y = \sin x$ is symmetric about the origin. That is, if the graph is rotated $180°$ about the origin, the result is the same as the original graph. What does this tell us about the relationship between $\sin x$ and $\sin(-x)$?

(b) What symmetry does the graph of $y = \cos x$ exhibit, and what does this symmetry tell us about the relationship between $\cos x$ and $\cos(-x)$?

Problem 2.17:

(a) What is the period of $\tan x$?

(b) Graph $y = \tan x$.

(c) We call a line an **asymptote** of a graph if the graph gets closer and closer to the line, but does not intersect it, as x or y gets far from 0. What lines are asymptotes of the graph of $y = \tan x$?

Problem 2.18:

(a) Graph $y = \csc x$.

(b) What is the period of $\csc x$?

(c) What lines are asymptotes of $\csc x$?

(d) At what points do the graphs of $y = \sin x$ and $y = \csc x$ intersect?

(e) At what points do the graphs of $y = \cos x$ and $y = \csc x$ intersect?

Problem 2.19: What is the number of solutions of the equation $\frac{x}{100} = \sin x$? *(Source: AMC)*

Problem 2.13:

(a) Find the graph of $y = \sin x$ for $0 \le x \le 2\pi$.

(b) Use your answer to part (a) to construct the graph of $y = \sin x$ for all x.

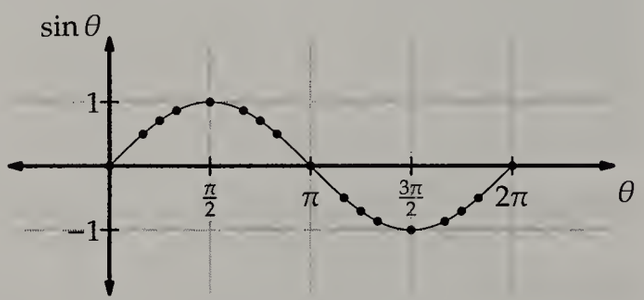
(c) A function f is called **periodic** if there is some nonzero constant t such that $f(x) = f(x + t)$ for all values of x such that $f(x)$ and $f(x + t)$ are defined. The smallest possible positive value of t is called the **period** of f (if such a smallest value exists). Show that $g(x) = \sin x$ is periodic and find its period.

Solution for Problem 2.13:

(a) We'll produce the graph of $y = \sin x$ by choosing a variety of angles θ and evaluating $\sin\theta$ for those angles. We'll then plot the points $(\theta, \sin\theta)$, and connect them to build our graph. We use θ here rather than x because we want to avoid confusion when speaking of x-coordinates of points on the unit circle.

We start by choosing values of θ for which we can compute $\sin\theta$ easily. We build a table as shown below, and use the table to shape the corresponding graph, which is shown at right below.

θ	$\sin\theta$	θ	$\sin\theta$
0	0	π	0
$\pi/6$	1/2	$7\pi/6$	$-1/2$
$\pi/4$	$\sqrt{2}/2$	$5\pi/4$	$-\sqrt{2}/2$
$\pi/3$	$\sqrt{3}/2$	$4\pi/3$	$-\sqrt{3}/2$
$\pi/2$	1	$3\pi/2$	-1
$2\pi/3$	$\sqrt{3}/2$	$5\pi/3$	$-\sqrt{3}/2$
$3\pi/4$	$\sqrt{2}/2$	$7\pi/4$	$-\sqrt{2}/2$
$5\pi/6$	1/2	$11\pi/6$	$-1/2$
		2π	0



We can also get an intuition for the shape of the graph by thinking about the unit circle. The y-coordinate of the terminal point of θ is $\sin\theta$, so we track how the y-coordinate of the terminal point of θ changes as θ ranges from 0 to 2π. As θ increases from 0 to $\pi/2$, the y-coordinate of the terminal point of θ increases from 0 to 1. Moreover, this y-coordinate increases more quickly when θ is near 0 than when θ is near $\pi/2$. This corresponds to the way the graph above increases quickly near $\theta = 0$, but the pace of this increasing slows as θ nears $\pi/2$. Continuing past $\theta = \pi/2$, the y-coordinate of the terminal point decreases from 1 to 0 as θ increases from $\pi/2$ to π. This corresponds to the downward trajectory of our graph from $\theta = \pi/2$ to $\theta = \pi$.

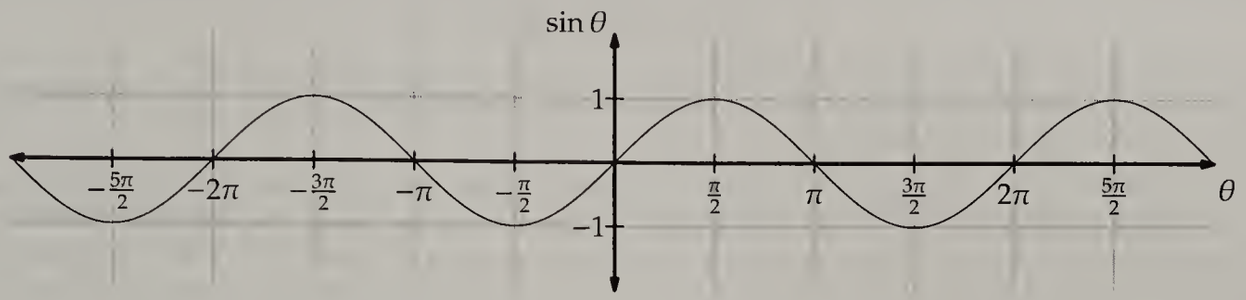
Then, θ continues from π to $3\pi/2$, and the y-coordinate of its terminal point goes from 0 to -1. Finally, θ goes from $3\pi/2$ to 2π, during which time the y-coordinate of its terminal point goes from -1 to 0. So, as $\sin\theta$ goes from 0 to -1 and back to 0 while θ goes from π to 2π, we get the curve below the horizontal axis in the graph of above.

(b) We find the rest of the graph by continuing with the unit circle. The value of $\sin\theta$ is the y-coordinate of the terminal point of the angle θ on the unit circle. The angle θ has the same terminal point as the angle $\theta + 2\pi$, and as the angle $\theta + 4\pi$, and as the angle $\theta + 6\pi$, and so on. Similarly, it has the same terminal point as the angles $\theta - 2\pi$, $\theta - 4\pi$, and so on. So, we have

$$\cdots = \sin(\theta - 4\pi) = \sin(\theta - 2\pi) = \sin\theta = \sin(\theta + 2\pi) = \sin(\theta + 4\pi) = \cdots.$$

In other words, the value of $\sin\theta$ stays the same whenever we increase or decrease θ by 2π. This means the vertical coordinate of a point on the graph stays the same each time we increase or decrease θ by 2π. In

other words, the graph repeats itself over and over in both directions forever, giving us the graph below.



We can now replace "θ" and "$\sin\theta$" in the graph above to produce the graph of $y = \sin x$. We will also typically remove the 1 and -1 on the vertical axis to reduce the clutter of the graph, as shown below.

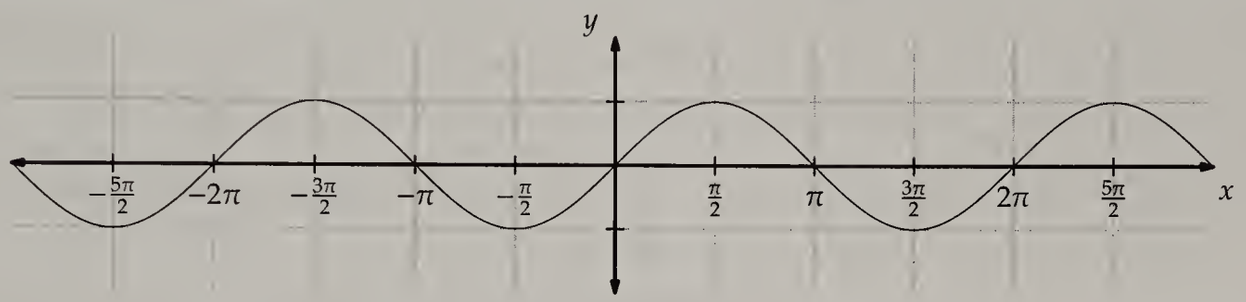


Figure 2.5: Graph of $y = \sin x$

The only difference between this and our previous graph is the labeling of the axes, where we labeled the axes of our first graph with θ and $\sin\theta$ to allow use to use "x" and "y" in discussing the unit circle in our solution.

(c) As we explained in the previous part, we have

$$\sin x = \sin(x + 2\pi)$$

for all values of x. Therefore, $g(x) = \sin x$ is indeed a periodic function, and the fact that $\sin x = \sin(x + 2\pi)$ for all x tells us that its period is no more than 2π. We can see from the graph that the period is no smaller than 2π, but to show this algebraically, we consider the points at which $\sin x = 1$. The only angles x for which $\sin x = 1$ are those with terminal point $(0, 1)$. The smallest such positive angle is $x = \frac{\pi}{2}$ and the next smallest occurs a full revolution around the unit circle later, when $x = \frac{5\pi}{2}$. These two angles for which $\sin x = 1$ are 2π apart, and there are no angles x between $\frac{\pi}{2}$ and $\frac{5\pi}{2}$ such that $\sin x = 1$. Therefore, the period of $\sin x$ cannot be less than 2π.

We've now shown that $\sin x = \sin(x + 2\pi)$ for all x, and that there is no positive value of t smaller than $t = 2\pi$ such that $\sin x = \sin(x + t)$ for all x, so we conclude that the period of $g(x) = \sin x$ is 2π.

If a function is periodic with period t, we sometimes refer to any interval of length t as a "period" of the function. So, for example, $[0, 2\pi]$ is one period of $\sin x$, as are $[2\pi, 4\pi]$, $[6, 6 + 2\pi]$, and any other interval of length 2π.

□

Problem 2.14: Graph $y = \cos x$. What is the period of $\cos x$?

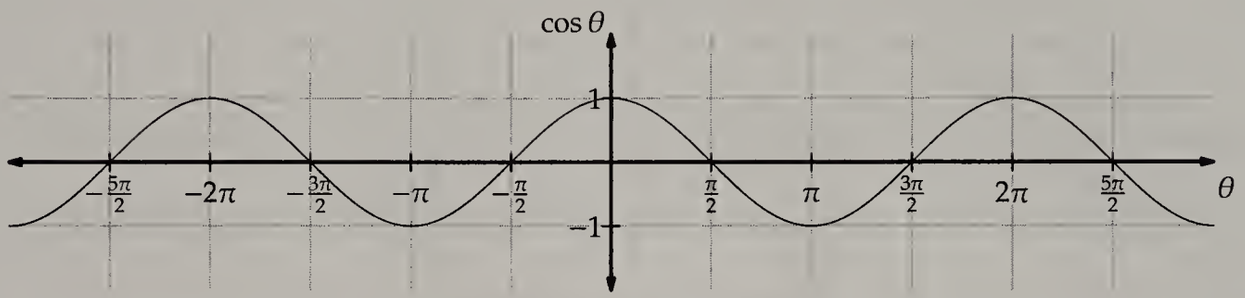
Solution for Problem 2.14: Just as we did with the graph of $\sin x$, we will consider the unit circle to develop the graph of $y = \cos x$. As before, we'll use θ as the angle, to avoid confusion as we refer to the unit circle.

For starters, we can consider the unit circle to see that

$$\cos\theta = \cos(\theta + 2\pi)$$

for all values of θ, because the terminal point of θ is the same as the terminal point of $\theta + 2\pi$. Therefore, the graph of $\cos\theta$ repeats itself horizontally every 2π.

To see how the graph behaves from $\theta = 0$ to $\theta = 2\pi$, we consider the unit circle. When $\theta = 0$, the terminal point of θ is $(1,0)$, so $\cos\theta = 1$ for $\theta = 0$. As θ increases from 0 to $\frac{\pi}{2}$, the value of $\cos\theta$ decreases from 1 to 0, which corresponds to the fact that the x-coordinate of the terminal point of θ goes from 1 to 0 as θ goes from 0 to $\frac{\pi}{2}$. Continuing around the unit circle, as θ goes from $\frac{\pi}{2}$ to π, the terminal point of θ goes from $(0,1)$ to $(-1,0)$. The x-coordinate, and therefore the value of $\cos\theta$, goes from 0 to -1. This is illustrated in the graph as the graph goes below the horizontal axis when θ increases past $\frac{\pi}{2}$. The graph reaches its lowest point at $\theta = \pi$, where $\cos\theta = -1$, and then goes upward as θ increases past π. The resulting graph is shown below.



Back on the unit circle, as θ goes from π to $\frac{3\pi}{2}$, the x-coordinate (which is $\cos\theta$) of the terminal point goes from -1 to 0. Correspondingly, the graph rises as θ goes from π to $\frac{3\pi}{2}$. The graph intersects the horizontal axis at $\theta = \frac{3\pi}{2}$. It then continues upward, reflecting the fact that the x-coordinate of a fourth quadrant angle's terminal point is positive. The graph again reaches $\cos\theta = 1$ when we complete our revolution about the the unit circle, which is when $\theta = 2\pi$. We're then back where we started on the unit circle, so the graph as θ increases from 2π is the same as the graph is as θ increases from 0. Therefore, cosine is periodic with period 2π.

We can replace the horizontal and vertical labels of our graph above to give the graph of $y = \cos x$:

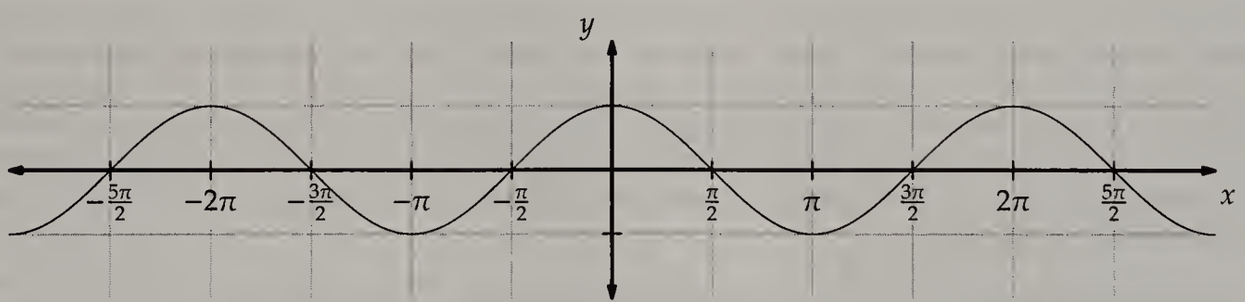


Figure 2.6: Graph of $y = \cos x$

□

The shape of the graphs of sine and cosine is sometimes referred to as a **sinusoid** or **sine curve**. You might have noticed that these two graphs look *very* similar.

Problem 2.15: What is the relationship between the graph of $y = \sin x$ and the graph of $y = \cos x$? What does this suggest as a relationship between sine and cosine?

Solution for Problem 2.15: We graph both $y = \sin x$ (solid) and $y = \cos x$ (dashed) on the same Cartesian plane below.

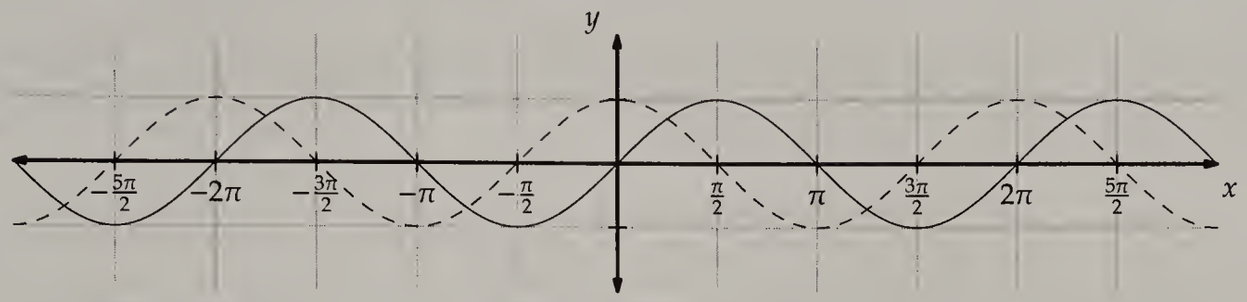


Figure 2.7: Graphs of $y = \sin x$ (solid) and $y = \cos x$ (dashed)

The graph of $y = \cos x$ appears to be a leftward shift of the graph of $y = \sin x$. This suggests an interesting relationship between sine and cosine. Specifically, we let $f(x) = \sin x$. In Section 1.2, we saw that shifting the graph of $y = f(x)$ leftward by $\frac{\pi}{2}$ gives us the graph of $y = f\left(x + \frac{\pi}{2}\right)$. Our graphs above suggest that this shift produces the graph of $y = \cos x$. So, it appears that $f\left(x + \frac{\pi}{2}\right) = \cos x$ for all values of x. That is, it looks like we have a new trigonometric identity:

$$\sin\left(x + \frac{\pi}{2}\right) = \cos x.$$

We'll learn how to prove this identity, and many more, in Chapter 3. □

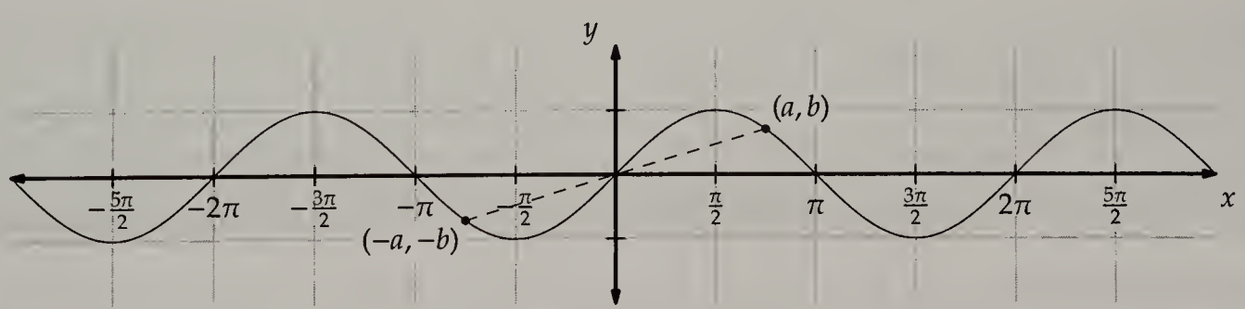
Problem 2.16:

(a) The graph of $y = \sin x$ is symmetric about the origin. That is, if the graph is rotated 180° about the origin, the result is the same as the original graph. What does this tell us about the relationship between $\sin x$ and $\sin(-x)$?

(b) What symmetry does the graph of $y = \cos x$ exhibit, and what does this symmetry tell us about the relationship between $\cos x$ and $\cos(-x)$?

Solution for Problem 2.16:

(a) If a point P is rotated 180° about point O to produce point P', then O is the midpoint of $\overline{PP'}$. Therefore, if (a, b) is rotated 180° about $(0, 0)$, the result must be $(-a, -b)$, since $(0, 0)$ is the midpoint of the segment with endpoints (a, b) and $(-a, -b)$.

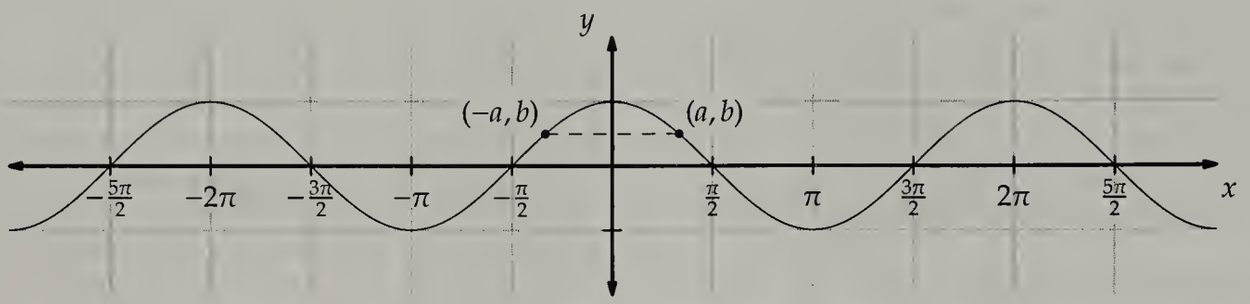


Since the graph of $y = \sin x$ is symmetric about the origin, we know that for every point (a, b) on the graph of $y = \sin x$, the point $(-a, -b)$ is also on the graph of $y = \sin x$. Because (a, b) and $(-a, -b)$ are on the graph of $y = \sin x$, we have $b = \sin a$ and $-b = \sin(-a)$. Putting these together gives us $\sin(-a) = -b = -\sin a$. Therefore, the fact that the graph of $y = \sin x$ is symmetric about the origin tells us that for all x, we have

$$\sin(-x) = -\sin x.$$

A function f such that $f(-x) = -f(x)$ for all x in the domain of f is called an **odd function**. Therefore, sine is an odd function.

(b) The graph of $y = \cos x$ is symmetric about the y-axis. This means that for every point (a, b) on the graph, its mirror image over the y-axis, which is $(-a, b)$, is also on the graph. Therefore, we have both $b = \cos a$ and $b = \cos(-a)$, so we have $\cos a = \cos(-a)$ for all a.



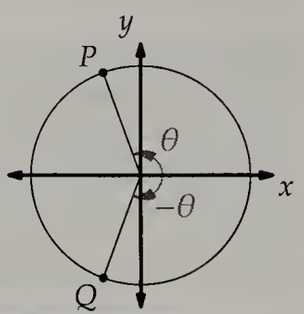
A function f such that $f(-x) = f(x)$ for all x in the domain of f is called an **even function**. Because

$$\cos(-x) = \cos x$$

for all x, cosine is an even function.

□

We can prove the relationships suggested in Problem 2.16 by using the unit circle. The actions of rotating the point $(1, 0)$ an angle θ counterclockwise and rotating $(1, 0)$ an angle θ clockwise are symmetric about the x-axis. That is, if P is the result of rotating $(1, 0)$ an angle θ counterclockwise about the origin and Q is the result of rotating $(1, 0)$ an angle θ clockwise about the origin, then P and Q are reflections of each other over the x-axis. Therefore, P and Q have the same x-coordinate, but opposite y-coordinates, which means that $\cos(-\theta) = \cos\theta$ and $\sin(-\theta) = -\sin\theta$.



Important:

$$\cos(-\theta) = \cos\theta,$$
$$\sin(-\theta) = -\sin\theta.$$

Problem 2.17:

(a) What is the period of $\tan x$?

(b) Graph $y = \tan x$.

(c) We call a line an **asymptote** of a graph if the graph gets closer and closer to the line, but does not intersect it, as x or y gets far from 0. What lines are asymptotes of the graph of $y = \tan x$?

Solution for Problem 2.17:

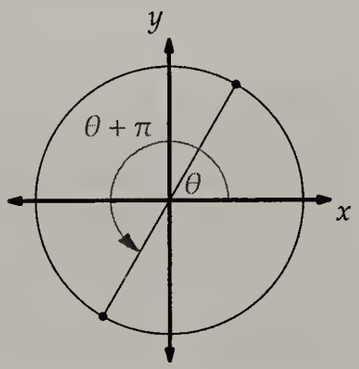
(a) The period of $\sin x$ is 2π and the period of $\cos x$ is 2π, so we might guess that the period of $\tan x$ is 2π as well. After all, since $\sin(x + 2\pi) = \sin x$ and $\cos(x + 2\pi) = \cos x$, we have

$$\tan(x + 2\pi) = \frac{\sin(x + 2\pi)}{\cos(x + 2\pi)} = \frac{\sin x}{\cos x} = \tan x.$$

While this does tell us that $\tan x$ is periodic, and that the period is no greater than 2π, it doesn't tell us that the period is exactly 2π, since the period could still be smaller than 2π.

We turn to the unit circle to determine the period of $\tan x$. We defined $\tan\theta$ as the slope of the line through the origin and the terminal point of θ. As θ ranges from 0 to $\frac{\pi}{2}$, this slope starts at 0 and increases without bound, without ever being the same for two different values of θ. Similarly, as θ goes from 0 down to $-\frac{\pi}{2}$, this slope starts at 0 and decreases without bound, again without ever being the same for two different values of θ. Moreover, the slope is 0 when $\theta = 0$, negative when $0 > \theta > -\frac{\pi}{2}$, and positive when $0 < \theta < \frac{\pi}{2}$, so there are no two values of θ from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ that have the same tangent. This means that the period of tangent cannot be less than π.

But can the period of tangent *equal* π? Again, the unit circle holds the answer. For any m, the line through the origin with slope m intersects the unit circle at two points that are diametrically opposite each other. Therefore, they are the terminal points of two angles that differ by π. Moreover, because the same line through the origin passes through both terminal points, the tangents of these two angles are the same. Therefore, we have $\tan(\theta + \pi) = \tan\theta$ for all θ for which $\tan\theta$ is defined, so the period of tangent is π.



(b) We did all the hard work in the previous part. Since $\tan x$ has period π, we can graph $y = \tan x$ for a single period, and then just copy it over and over to get the graph of $y = \tan x$. We'll choose to graph $\tan x$ first for $-\frac{\pi}{2} < x < \frac{\pi}{2}$, since $\tan x$ is defined for all values of x in this range.

As discussed in part (a), we have $\tan 0 = 0$, and $\tan x$ increases without bound as x goes from 0 to $\frac{\pi}{2}$ (but $\tan x$ is not defined for $x = \frac{\pi}{2}$, so there are no points on the graph with this value of x). So, as x goes from 0 to $\frac{\pi}{2}$, the graph of $y = \tan x$ starts at the origin and goes upward forever as x approaches $\frac{\pi}{2}$.

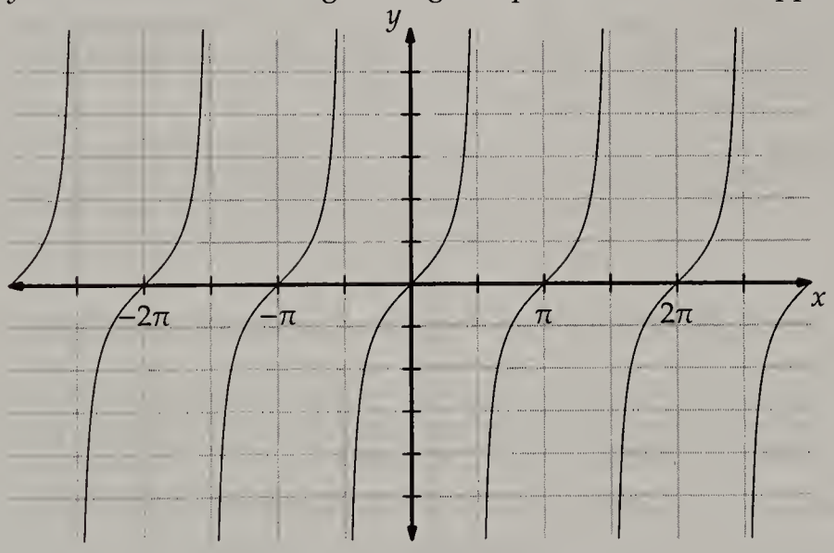


Figure 2.8: Graph of $y = \tan x$

The graph does not intersect the line $x = \frac{\pi}{2}$, because $\tan x$ is not defined when $x = \frac{\pi}{2}$. However, the graph does approach this line without ever reaching it. The line $x = \frac{\pi}{2}$ is called an **asymptote** of the graph of $y = \tan x$.

Moving over to the other side of the origin, we can follow the same reasoning as above to build the graph as x goes from 0 to $-\frac{\pi}{2}$. Since $\tan x$ decreases without bound as x goes from 0 to $-\frac{\pi}{2}$, the graph starts at $(0, 0)$ and goes downward forever as x approaches $-\frac{\pi}{2}$. Moreover, this portion of the graph is the 180° rotation about the origin of the portion for $0 < x < \frac{\pi}{2}$, because

$$\tan(-x) = \frac{\sin(-x)}{\cos(-x)} = \frac{-\sin x}{\cos x} = -\tan x.$$

Therefore, $\tan x$ is an odd function, and its graph is symmetric about the origin. Just as $x = \frac{\pi}{2}$ is an asymptote of the graph, so is $x = -\frac{\pi}{2}$.

Now that we have graphed one full period of $\tan x$, we can repeat this graph indefinitely in both directions to produce the graph of $y = \tan x$, as shown above.

(c) Since the graph repeats in intervals of π, its asymptotes repeat in intervals of π. We saw above that $x = \frac{\pi}{2}$ is an asymptote of the graph, so other asymptotes are $x = \frac{3\pi}{2}$, $x = \frac{5\pi}{2}$, $x = \frac{7\pi}{2}$, etc., as well as $x = -\frac{\pi}{2}$, $x = -\frac{3\pi}{2}$, $x = -\frac{5\pi}{2}$, etc. We can express these all at once: the asymptotes are $x = \frac{n\pi}{2}$ for all odd integers n. Note that these values correspond to the values of x for which $\cos x = 0$, which means $\tan x$ is not defined.

□

Problem 2.18:

(a) Graph $y = \csc x$.

(b) What is the period of $\csc x$?

(c) What lines are asymptotes of $\csc x$?

(d) At what points do the graphs of $y = \sin x$ and $y = \csc x$ intersect?

(e) At what points do the graphs of $y = \cos x$ and $y = \csc x$ intersect?

Solution for Problem 2.18:

(a) Since $\sin 0 = 0$, $\csc 0$ is undefined, so the graph of $y = \csc x$ does not intersect the y-axis. When x is a very small positive number, $\sin x$ is very small, so $\csc x$ is very large. Similarly, when x is the the negative of a very small number, $\sin x$ is negative, but very close to 0, which makes $\csc x$ the negative of a very large number. So, just to the right of the y-axis, the graph of $y = \csc x$ is very far above the x-axis, while just to the left of the y-axis, the graph is very far below the x-axis.

As x increases from 0 to $\frac{\pi}{2}$, $\sin x$ goes from 0 up to 1. As $\sin x$ increases, $\frac{1}{\sin x}$ decreases until it reaches 1 at $x = \frac{\pi}{2}$. Then, as x goes from $\frac{\pi}{2}$ up to π, $\sin x$ goes from 1 back down to 0, and $\csc x$ increases without bound, starting from $\csc x = 1$ at $x = \frac{\pi}{2}$. Just as the graph does not intersect the line $x = 0$, it does not intersect $x = \pi$.

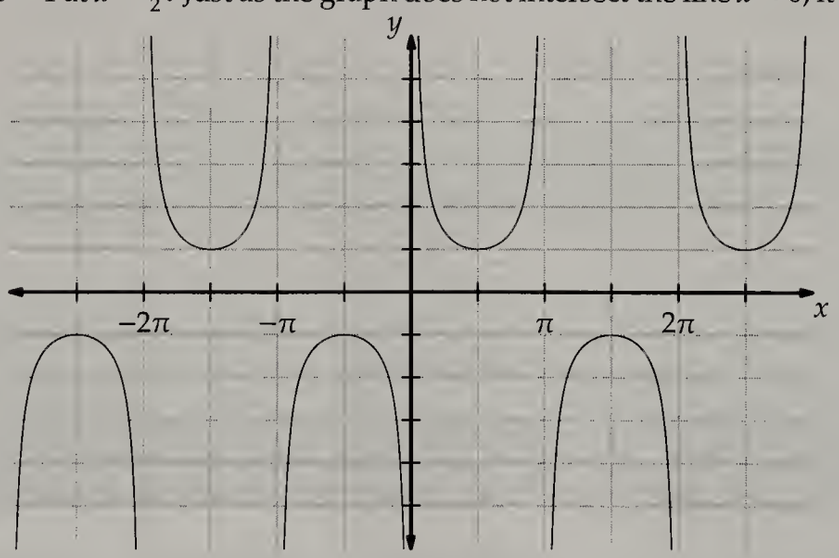


Figure 2.9: Graph of $y = \csc x$

On the other side of the y-axis, the pattern is repeated, but all the action is below the x-axis. As x goes from 0 to $-\frac{\pi}{2}$, $\sin x$ goes from 0 to -1, so the graph of $y = \csc x$ goes from very far below the x-axis, up to the point $\left(-\frac{\pi}{2}, -1\right)$. Then, as x goes from $-\frac{\pi}{2}$ to $-\pi$, $\sin x$ goes from -1 back to 0. So, the graph of $y = \csc x$ goes back down from $\left(-\frac{\pi}{2}, -1\right)$, continuing indefinitely below the x-axis as the graph approaches, but never reaches, the line $x = -\pi$.

But what happens beyond $x = \pi$ and $x = -\pi$? Fortunately, we don't have to try to figure that out from scratch. Since $\sin x$ is periodic with period 2π, we know that $\frac{1}{\sin x}$ must also repeat itself in intervals of 2π. Since we have already determined the graph from $x = -\pi$ to $x = \pi$, we can simply repeat this portion of the graph every 2π to produce the full graph of $y = \csc x$ shown earlier.

(b) As discussed in the previous part, since the period of $\sin x$ is 2π, we know that the period of $\frac{1}{\sin x}$ is also 2π.

(c) Since $\sin x = 0$ whenever x is an integer multiple of π, $\csc x$ is undefined when x is an integer multiple of π. However, x can take on values very close to these integer multiples of π. At these values of x, the graph of $y = \csc x$ is very far from the x-axis. We therefore see that the lines $x = n\pi$, where n is an integer, are asymptotes of the graph of $y = \csc x$.

(d) If the graphs of $y = \sin x$ and $y = \csc x$ intersect, then we must have $\sin x = \csc x$. We therefore must have $\sin x = 1/(\sin x)$, or $\sin^2 x = 1$, which means $\sin x = 1$ or $\sin x = -1$. This occurs when $x = \frac{n\pi}{2}$ for all odd integers n, which tells us that our graphs intersect at all these values of x. We can see this in the graph below, in which $y = \sin x$ is graphed as a dashed curve, and $y = \csc x$ is graphed as a solid curve.

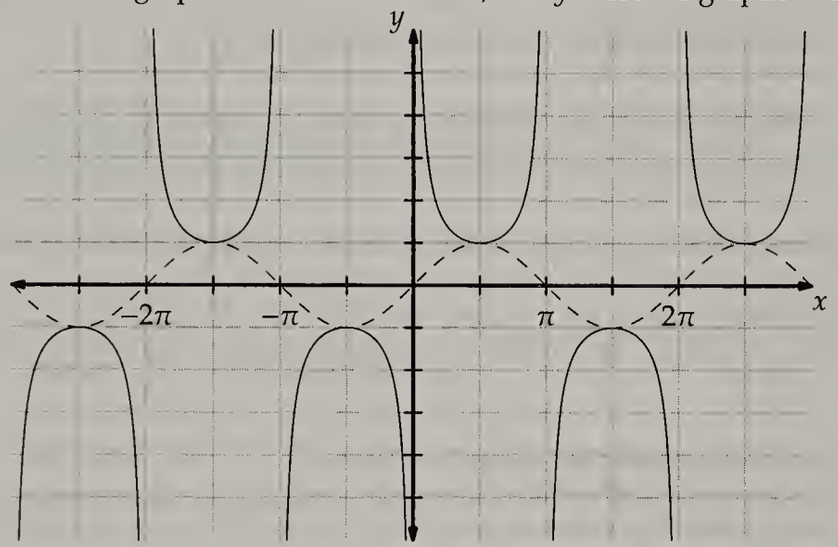


Figure 2.10: Graphs of $y = \csc x$ (solid) and $y = \sin x$ (dashed)

(e) If $y = \cos x$ and $y = \csc x$ intersect, then we must have $\cos x = \csc x$, so $\cos x = \frac{1}{\sin x}$. Therefore, we must have $(\cos x)(\sin x) = 1$. However, neither $\cos x$ nor $\sin x$ can exceed 1 or be less than -1. So, if $(\cos x)(\sin x) = 1$, then we must have $\cos x = \sin x = 1$ or $\cos x = \sin x = -1$. But both of these are impossible, since $\sin x = 0$ whenever $\cos x$ is -1 or 1. Therefore, the two graphs do not intersect.

We also could have seen that the graphs do not intersect by noting that the graph of $y = \csc x$ never goes between the lines $y = 1$ and $y = -1$, and the graph of $y = \cos x$ never goes outside them. Therefore, the graphs can only intersect on one of these lines. The graph of $y = \cos x$ only meets these lines when $x = n\pi$ for integers n, but the graph of $y = \csc x$ does not have any points for which x is an integer multiple of π, because $\csc x$ is undefined for these values of x. Therefore, the graphs do not intersect.

□

Problem 2.19: What is the number of solutions of the equation $\frac{x}{100} = \sin x$? *(Source: AMC)*

Solution for Problem 2.19: We can't isolate x or factor this equation in any useful way. We have one obvious solution at $x = 0$, which makes both sides equal to 0, but there aren't any other obvious solutions. Maybe there are no other solutions?

To get some intuition for the problem, we try graphing both sides of the equation. That is, we graph $y = \frac{x}{100}$ and $y = \sin x$. The intersection points of these graphs correspond to solutions of the equation $\frac{x}{100} = \sin x$.

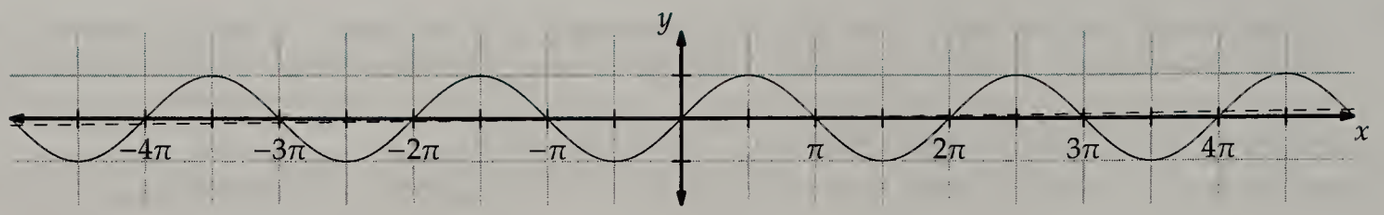


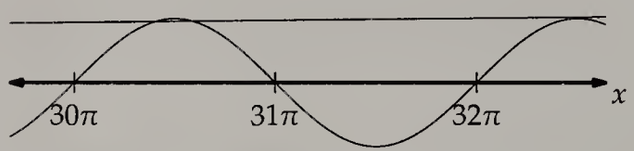
Figure 2.11: Graphs of $y = \sin x$ (solid) and $y = \frac{x}{100}$ (dashed)

Aha! There are many more solutions. If fact, we see that for each period of $\sin x$, a line that is between $y = 0$ and $y = 1$ will intersect the graph of $y = \sin x$ twice—once as $\sin x$ goes up from 0 to 1, and then again as $\sin x$ goes down from 1 to 0. The same is true for any line that is between $y = 0$ and $y = -1$ for a full period of $\sin x$.

To count the number of intersections, we first note that both graphs are symmetric about the origin. So, counting the intersections on the right of the y-axis gives us the number of intersections on the left of the y-axis, as well. We also have to remember the origin, and add that in at the end.

WARNING!! Always be on the lookout for special cases. 0 is a common special case.

In counting the intersections to the right of the y-axis, we note that the period from 0 to 2π has only 1 intersection with the line $y = \frac{x}{100}$. After that, each period of 2π has two intersections. For example, we can see that the dashed line intersects the sine curve just past $x = 2\pi$ and just before $x = 3\pi$. The line $y = \frac{x}{100}$ will exceed $y = 1$ when $x > 100$, after which the line will not intersect the graph of $y = \sin x$ anymore. So, now we just have to count the number of periods up to $x = 100$.



To count the number of periods up to $x = 100$, we note that $30\pi \approx 94.2$ and $32\pi \approx 100.48$. So, we know that we complete 15 full periods and then most of another one as x goes from 0 to 100. But how do we know if the line intersects the graph of $y = \sin x$ twice during the period from $x = 30\pi$ to $x = 32\pi$? The diagram above shows the portion of both graphs for values of x from a little less than 30π to a little more than 32π. This picture isn't terribly illuminating. It's not even obvious from the diagram that graphs don't intersect for $x > 32\pi$. We have to use our observation that $32\pi > 100$ to be sure that $\frac{x}{100} > \sin x$ for all $x > \frac{32\pi}{100}$, since $\frac{x}{100} > 1 \geq \sin x$ for $x > 32\pi$. Perhaps we can come up with a similar argument to see that the graphs intersect twice when $30\pi < x < 31\pi$.

We consider where in the period the line meets the sine curve. The sine curve is above the x-axis in the first half of each period (from $x = 0$ to $x = \pi$, from $x = 2\pi$ to $x = 3\pi$, etc), so it is above the x-axis from $x = 30\pi$ to $x = 31\pi$. Since the line is above the x-axis when $x > 0$, both of its intersections with $y = \sin x$ in the interval $30\pi < x < 32\pi$ must be in the range $30\pi < x < 31\pi$. Now, we see the line does indeed intersect this final period twice, because $31\pi \approx 97.34$.

There is one subtle point the previous paragraph overlooks. We have shown that if the graph of the line intersects the period of $\sin x$ from $x = 30\pi$ to $x = 32\pi$, then it does so between $x = 30\pi$ and $x = 31\pi$, but we haven't shown that it's impossible for the line to be *tangent* to the graph of $y = \sin x$, so that there is only one intersection point rather than two. To see why there are at least two intersection points, we note that $\sin 30.5\pi = 1$, and $(30.5\pi)/100 \approx 0.9582$, which is less than 1. Since $\sin 30\pi = \sin 31\pi = 0$, we see that the graph of $y = \frac{x}{100}$ is above the graph of $y = \sin x$ when $x = 30\pi$, below it when $x = 30.5\pi$, and above it again when $x = 31\pi$. Therefore, the line intersects the graph of $y = \sin x$ at more than 1 point for $30\pi < x < 32\pi$. (We need more advanced tools to give a rigorous explanation why the line doesn't intersect $y = \sin x$ at more than two points for $30\pi < x < 32\pi$.)

We're now ready to count all the intersections. To the right of the y-axis, the line intersects the sine curve 1 time in the first period, and 2 times in each of the next 15 periods, for a total of $1 + 2 \cdot 15 = 31$ intersections. Both graphs

are symmetric about the origin, so they intersect 31 times to the left of the y-axis, too. Adding in the intersection at the origin gives a total of $31 + 31 + 1 = 63$ intersections, so there are 63 solutions to the equation $\frac{x}{100} = \sin x$. □

Concept: Graphs can help provide intuition for problems involving trig functions.

Exercises

2.3.1

(a) Graph $y = \sec x$.

(b) What is the period of $\sec x$?

(c) What is the relationship between the graph of $y = \sec x$ and the graph of $y = \csc x$?

2.3.2

(a) Graph $y = \cot x$.

(b) What is the period of $\cot x$?

2.3.3 At how many points with x-coordinate between 0 and 100 do the graphs of $y = \sin x$ and $y = \cos x$ intersect?

2.4 Transformations of Trig Functions

For the graphing problems in this section, try using the methods we discussed in Section 1.2 for transforming graphs of functions. After producing each graph, test your answer by trying various values of x, finding the corresponding value of y, and checking that the point (x, y) is indeed on the graph you produced.

Problems

Problem 2.20: Graph $y = 2\sin x$ and $y = \sin x + 2$.

Problem 2.21:

(a) Graph $y = \cos 2x$.

(b) What is the period of $\cos 2x$?

(c) What is the period of $\cos ax$, where a is a positive constant?

Problem 2.22: What is the period of $\sin\left(-\frac{x}{3}\right)$?

Problem 2.23: Is $f(x) = \sin\frac{3x}{2} + \cos 2x$ periodic? If so, what is the period?

Problem 2.24: The **phase shift** of a function of the form $f(x) = a\sin(bx + c)$, where a, b, and c are constants, is the smallest horizontal distance by which the graph of $g(x) = a\sin bx$ must be shifted to produce the graph of f. We typically use positive values for the phase shift if this shift is to the right, and negative if the shift is to the left.

(a) What is the phase shift of $\sin\left(x - \frac{\pi}{2}\right)$?

(b) Graph $y = \sin\left(x - \frac{\pi}{2}\right)$.

Problem 2.25: Graph $y = \sin x + \sin(x - \pi)$. What does the graph suggest about the relationship between $\sin x$ and $\sin(x - \pi)$?

Problem 2.26:

(a) What is the phase shift of $-2\sin\left(3x - \frac{\pi}{4}\right)$? (Note: The answer is *not* $\frac{\pi}{4}$!) **Hints:** 62

(b) What is the period of $-2\sin\left(3x - \frac{\pi}{4}\right)$?

(c) Graph $y = -2\sin\left(3x - \frac{\pi}{4}\right)$.

Problem 2.27: A Ferris wheel is a ride that consists of seats connected to the circumference of a large vertical wheel. As the wheel rotates, the riders are moved from the bottom of the wheel to the top of the wheel, and then back to the bottom. When you first get on a particular Ferris wheel, you are at the bottom of the ride, 3 feet off the ground. Once the Ferris wheel starts, it rotates around its center at a constant rate. Suppose that it takes you 8 seconds to reach the highest point of the ride, 63 feet off the ground.

(a) You start at the bottom of the ride. How high off the ground will you be 5 seconds after the ride starts? (Answer to the nearest tenth of a foot.)

(b) Find a function f such that $f(t)$ equals how high in feet you will be off the ground t seconds later.

Problem 2.20: Graph $y = 2\sin x$ and $y = \sin x + 2$.

Solution for Problem 2.20: We can produce the graph of $y = 2\sin x$ by scaling the graph of $y = \sin x$ vertically away from the x-axis by a factor of 2. The resulting graph is shown below:

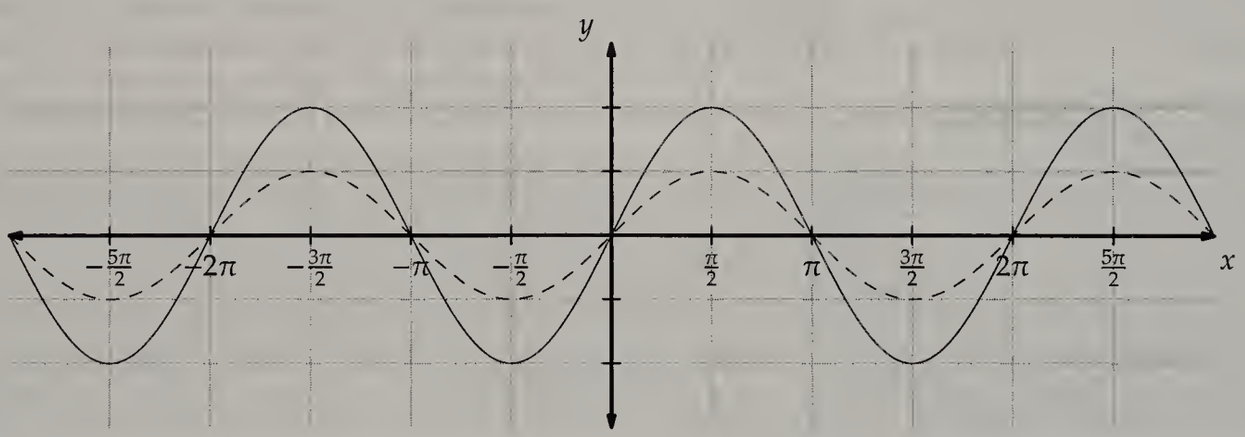


Figure 2.12: Graphs of $y = 2\sin x$ (solid) and $y = \sin x$ (dashed)

The only difference between the graph of $y = 2\sin x$ and the graph of $y = \sin x$ is the vertical scale. We use **amplitude** as a measure of this vertical scale, where the amplitude equals half the vertical distance between a maximal (highest) and minimal (lowest) point of the graph. Equivalently, we can say that the amplitude of a sine (or cosine) function is half the difference between the maximum and minimum values of the function. So, the amplitude of $\sin x$ is 1, and the amplitude of $2\sin x$ is 2.

We produce the graph of $y = \sin x + 2$ simply by shifting the graph of $y = \sin x$ up by 2 units, as shown atop the following page.

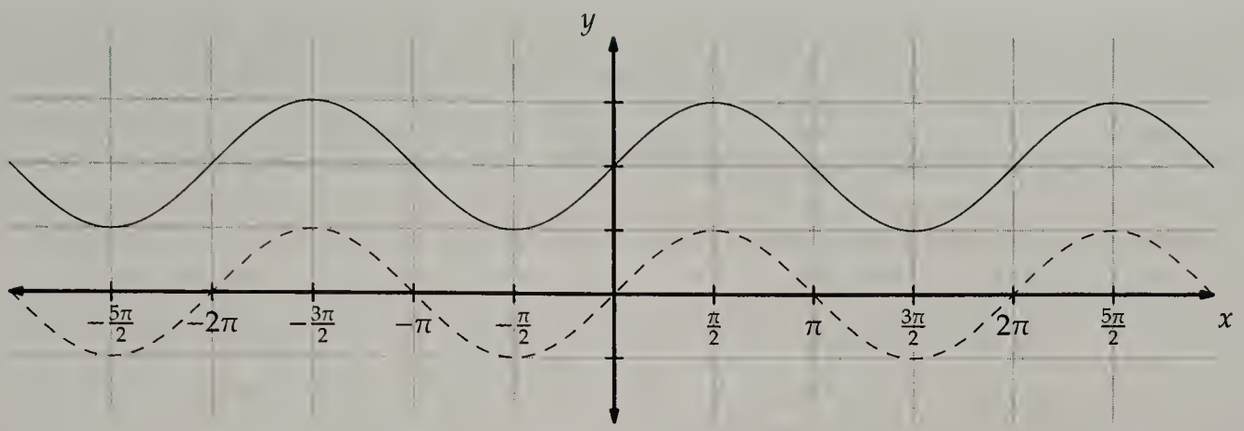


Figure 2.13: Graphs of $y = \sin x + 2$ (solid) and $y = \sin x$ (dashed)

Notice that shifting the graph vertically does not affect the vertical distance between a maximal (highest) and minimal (lowest) point on the graph. Therefore, the amplitude of the graph of $y = \sin x + 2$ is the same as the amplitude of the graph of $y = \sin x$, which is 1. □

As an Exercise, you'll show that the amplitude of $a \sin x$, where a is a constant, is $|a|$.

Problem 2.21:

(a) Graph $y = \cos 2x$.

(b) What is the period of $\cos 2x$?

(c) What is the period of $\cos ax$, where a is a positive constant?

Solution for Problem 2.21:

(a) In Section 1.2, we learned that the graph of $y = f(2x)$ is a horizontal scaling of the graph of $y = f(x)$ towards the y-axis by a factor of 1/2. So, to produce the graph of $y = \cos 2x$, we scale the graph of $y = \cos x$ towards the y-axis by a factor of 1/2, producing the graph below (the scaling of three specific points on the graph of $y = \cos x$ is indicated by the grey arrows):

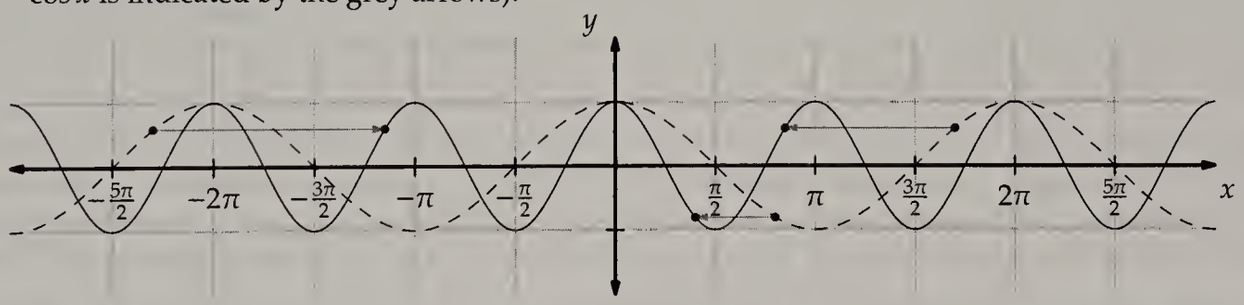


Figure 2.14: Graphs of $y = \cos 2x$ (solid) and $y = \cos x$ (dashed)

(b) The graph of $y = \cos x$ repeats every 2π. Since we compress this graph horizontally by a factor of 1/2 to produce the graph of $y = \cos 2x$, the graph of $y = \cos 2x$ repeats every $2\pi(1/2) = \pi$. Therefore, the period of $\cos 2x$ is π.

(c) The graph of $y = \cos ax$ results from compressing the graph of $y = \cos x$ horizontally by a factor of $1/a$. Since the graph of $y = \cos x$ repeats every 2π, the compressed graph of $y = \cos ax$ will repeat every $2\pi/a$. Therefore, the period of $\cos ax$ is $2\pi/a$.

□

Problem 2.22: What is the period of $\sin\left(-\frac{x}{3}\right)$?

Solution for Problem 2.22: We first deal with the negative sign. Since the graph of $y = f(-x)$ is the reflection of the graph of $y = f(x)$ over the y-axis (see page 14), the graph of $y = \sin(-x)$ is the reflection of the graph of $y = \sin x$ over the y-axis. Reflecting the graph of a periodic function over the y-axis doesn't have any effect on how frequently the graph repeats, so the period of $\sin(-x)$ is the same as the period of $\sin x$, which is 2π.

Next, we deal with the fraction. The graph of $y = f(x/3)$ is a horizontal stretching by a factor of $1/(1/3) = 3$ *away* from the x-axis. Therefore, the graph of $y = \sin\left(-\frac{x}{3}\right)$ results from stretching the graph of $y = \sin(-x)$ horizontally by a factor of 3. Since the graph of $y = \sin(-x)$ repeats every 2π, the graph of $y = \sin\left(-\frac{x}{3}\right)$ repeats every $3(2\pi) = 6\pi$. Therefore, $\sin\left(-\frac{x}{3}\right)$ has period 6π. □

We can follow the same reasoning as we did in Problems 2.21 and 2.22 to see that the period of both $\sin ax$ and $\cos ax$ is $\frac{2\pi}{|a|}$, where a is a nonzero constant.

Problem 2.23: Is $f(x) = \sin\frac{3x}{2} + \cos 2x$ periodic? If so, what is the period?

Solution for Problem 2.23: We'll start with the graph of $y = \sin\frac{3x}{2} + \cos 2x$, and check if the graph repeats itself regularly. Of course, you wouldn't want to build such a graph by hand! You'd use a computer, just like we did to produce the graph below:

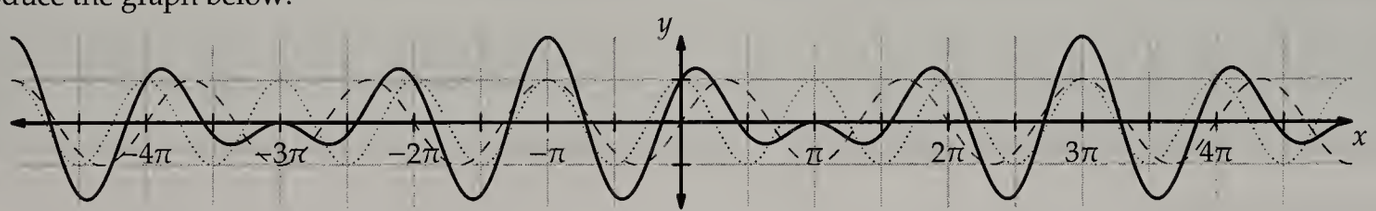


Figure 2.15: Graphs of $y = \sin\frac{3x}{2} + \cos 2x$ (solid), $y = \sin\frac{3x}{2}$ (dashed), and $y = \cos 2x$ (dotted)

The solid graph appears to peak at $x = -5\pi$, $x = -\pi$, and $x = 3\pi$. Each peak is 4π after the previous peak. The pattern of the solid curve from $x = -5\pi$ to $x = -\pi$ appears to be exactly the same as the curve from $x = -\pi$ to $x = 3\pi$, and the pattern appears to continue after $x = 3\pi$. It looks like the graph repeats itself, so we suspect the function $f(x) = \sin\frac{3x}{2} + \cos 2x$ is periodic. But how can we be sure? And if it is periodic, why is it periodic?

The large peaks give us one hint: these appear to occur where both $\sin\frac{3x}{2}$ and $\cos 2x$ equal 1. So, we investigate $\sin\frac{3x}{2}$ and $\cos 2x$ separately, with special attention on what values of x make either equal to 1. Both equal 1 at $x = -\pi$. As x increases, $\sin\frac{3x}{2}$ again equals 1 after it has gone through a full period. One period of $\sin\frac{3x}{2}$ has length $2\pi/(3/2) = \frac{4\pi}{3}$, so $\sin\frac{3x}{2}$ equals 1 at $x = -\pi + \frac{4\pi}{3} = \frac{\pi}{3}$, and again at $x = \frac{\pi}{3} + \frac{4\pi}{3} = \frac{5\pi}{3}$, and again at $x = \frac{5\pi}{3} + \frac{4\pi}{3} = 3\pi$.

Meanwhile, $\cos 2x$ has period $(2\pi)/2 = \pi$, so $\cos 2x$ equals 1 at $x = 0$, π, 2π, 3π, and so on. Indeed, both $\sin\frac{3x}{2}$ and $\cos 2x$ equal 1 again at 3π, and now we have an explanation for why they will both equal 1 again at $x = 3\pi + 4\pi = 7\pi$. As x increases by 4π, $\sin\frac{3x}{2}$ goes through exactly 3 full periods, while $\cos 2x$ goes through exactly 4 full periods. So, increasing x by 4π leaves both $\sin\frac{3x}{2}$ and $\cos 2x$ unchanged, which means that for $f(x) = \sin\frac{3x}{2} + \cos x$, we have $f(x + 4\pi) = f(x)$ for all x, so f is periodic. To see that the period is not smaller than 4π, we note that the values of x at which $f(x) = 2$ occur at intervals of 4π. □

The **phase shift** of a function of the form $f(x) = a\sin(bx + c)$, where a, b, and c are constants, is the smallest horizontal distance by which the graph of $g(x) = a\sin bx$ must be shifted to produce the graph of f. We typically use positive values for the phase shift if this shift is to the right, and negative if the shift is to the left.

Problem 2.24:
(a) What is the phase shift of $\sin\left(x - \frac{\pi}{2}\right)$?

(b) Graph $y = \sin\left(x - \frac{\pi}{2}\right)$.

Solution for Problem 2.24:

(a) As explained on page 13, the graph of $y = f(x - c)$ is a c-unit rightward shift of the graph of $y = f(x)$. Therefore, the graph of $y = \sin\left(x - \frac{\pi}{2}\right)$ is a $\frac{\pi}{2}$ rightward shift of the graph of $y = \sin x$, so the phase shift is $\frac{\pi}{2}$.

Another way to determine the phase shift is to compare the values of x such that $\sin x = 1$ to those at which $\sin\left(x - \frac{\pi}{2}\right) = 1$. We have $\sin x = 1$ for $x = \frac{\pi}{2}$, and at intervals of 2π thereafter. We have $\sin\left(x - \frac{\pi}{2}\right) = 1$ for $x = \pi$, and at intervals of 2π thereafter. So, the points at which $\sin\left(x - \frac{\pi}{2}\right) = 1$ are $\frac{\pi}{2}$ to the right of the points at which $\sin x = 1$.

(b) Since we found in part (a) that the graph of $y = \sin\left(x - \frac{\pi}{2}\right)$ is $\frac{\pi}{2}$ rightward shift of $\sin x$, we graph $y = \sin\left(x - \frac{\pi}{2}\right)$ by shifting the graph of $y = \sin x$ by $\frac{\pi}{2}$ to the right, as shown below:

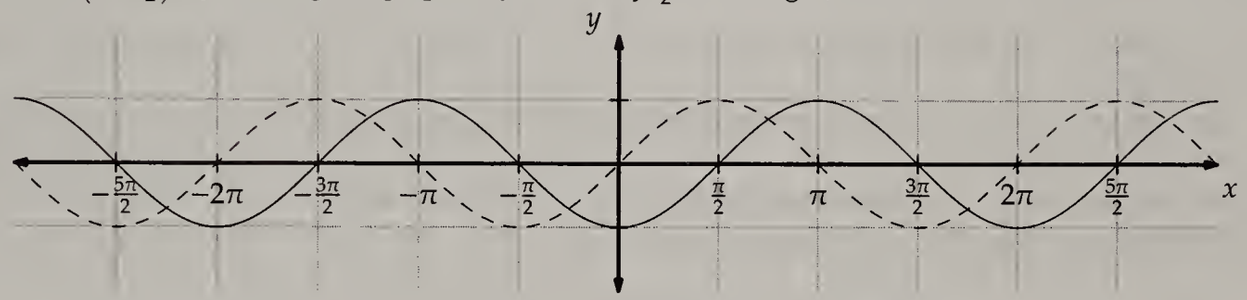


Figure 2.16: Graphs of $y = \sin\left(x - \frac{\pi}{2}\right)$ (solid) and $y = \sin x$ (dashed)

□

Problem 2.25: Graph $y = \sin x + \sin(x - \pi)$. What does the graph suggest about the relationship between $\sin x$ and $\sin(x - \pi)$?

Solution for Problem 2.25: The graph of $y = \sin(x - \pi)$ is a π-unit rightward shift of the graph of $y = \sin x$. We graph $y = \sin x$ and $y = \sin(x - \pi)$ below:

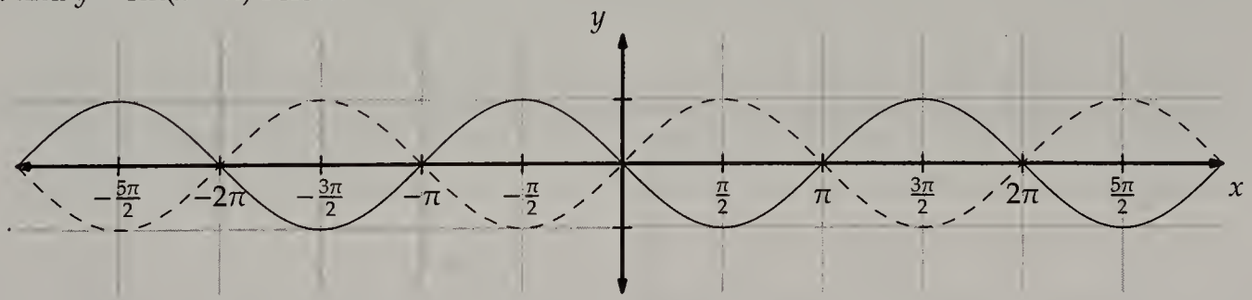


Figure 2.17: Graphs of $y = \sin(x - \pi)$ (solid) and $y = \sin x$ (dashed)

The graph of $y = \sin(x - \pi)$ appears to be the reflection of the graph of $y = \sin x$ over the x-axis. When we add the y-coordinates of the points on the graphs for each value of x, we always get 0, so the graph of $y = \sin x + \sin(x - \pi)$ is just the x-axis. This suggests that we have $\sin x + \sin(x - \pi) = 0$ for all values of x. That is, we appear to have another identity: $\sin(x - \pi) = -\sin x$. We will learn how to prove this identity algebraically in Chapter 3. □

Problem 2.26:

(a) What is the phase shift of $-2\sin\left(3x-\frac{\pi}{4}\right)$?

(b) What is the period of $-2\sin\left(3x-\frac{\pi}{4}\right)$?

(c) Graph $y=-2\sin\left(3x-\frac{\pi}{4}\right)$.

Solution for Problem 2.26:

(a) What's wrong with this solution:

Bogus Solution: The graph of $y=-2\sin\left(3x-\frac{\pi}{4}\right)$ is a $\frac{\pi}{4}$ rightward shift of the graph of $y=-2\sin 3x$, so the phase shift is $\frac{\pi}{4}$.

That there's a mistake here is clear when we consider where the graphs of $y=-2\sin\left(3x-\frac{\pi}{4}\right)$ and $y=-2\sin 3x$ cross the x-axis. Since $\sin 0=0$, the graph of $y=-2\sin 3x$ crosses the x-axis at $x=0$, while the graph of $y=-2\sin\left(3x-\frac{\pi}{4}\right)$ crosses the x-axis where $3x-\frac{\pi}{4}=0$, which is at $x=\frac{\pi}{12}$. Similarly, each x-intercept of the graph of $y=-2\sin\left(3x-\frac{\pi}{4}\right)$ is $\frac{\pi}{12}$ units to the right of an x-intercept of the graph of $y=-2\sin 3x$. So, this suggests that the graph of $y=-2\sin\left(3x-\frac{\pi}{4}\right)$ is a $\frac{\pi}{12}$ unit rightward shift of $y=-2\sin 3x$.

The heart of the error in the Bogus Solution is incorrectly applying the lessons of page 13 regarding horizontal shifts of graphs of functions. We learned there that the graph of $y=f(x-c)$ is a c-unit rightward shift of the graph of $y=f(x)$. However, if we let $f(x)=-2\sin 3x$, then $f\left(x-\frac{\pi}{4}\right)=-2\sin\left(3x-\frac{3\pi}{4}\right)$, which is *not* the same as $-2\sin\left(3x-\frac{\pi}{4}\right)$. Instead, we note that

$$-2\sin\left(3x-\frac{\pi}{4}\right)=-2\sin\left(3\left(x-\frac{\pi}{12}\right)\right)=f\left(x-\frac{\pi}{12}\right),$$

so the graph of $y=-2\sin\left(3x-\frac{\pi}{4}\right)$ is indeed a $\frac{\pi}{12}$ rightward shift of the graph of $y=-2\sin 3x$, which means that the phase shift is $\frac{\pi}{12}$.

(b) As explained in the previous part, the graph of $y=-2\sin\left(3x-\frac{\pi}{4}\right)$ is a rightward shift of the graph of $y=-2\sin 3x$. Translating a sine curve has no effect on its period, so the period of $-2\sin\left(3x-\frac{\pi}{4}\right)$ is the same as the period of $-2\sin 3x$. Similarly, the coefficient -2 in $-2\sin 3x$ has no impact on the period; it merely scales the graph vertically, which doesn't affect how frequently the graph repeats itself horizontally. So, the period of $-2\sin 3x$ is the same as the period of $\sin 3x$, which is $2\pi/3$. Therefore, the period of $-2\sin\left(3x-\frac{\pi}{4}\right)$ is $2\pi/3$.

(c) We can retrace our steps in the previous two parts to produce the graph. We start with the graph of $y=\sin 3x$, which is the graph of $y=\sin x$ compressed horizontally towards the y-axis by a factor of 3:

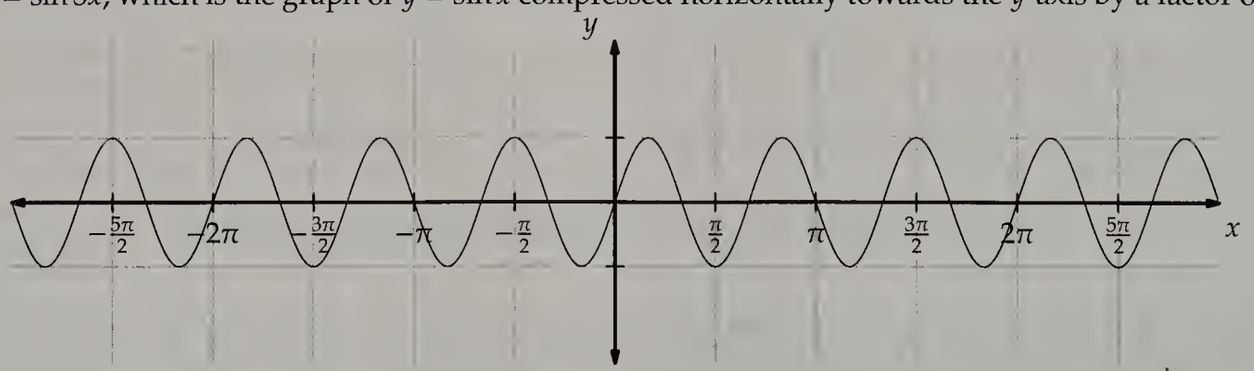


Figure 2.18: Graph $y=\sin 3x$

The graph of $y = 2\sin 3x$ is the result of scaling the graph of $y = \sin 3x$ vertically away from the x-axis by a factor of 2. The graph of $y = -2\sin 3x$ is the result of reflecting the graph of $y = 2\sin 3x$ over the x-axis (see page 14 if you don't see why), which gives us the graph below:

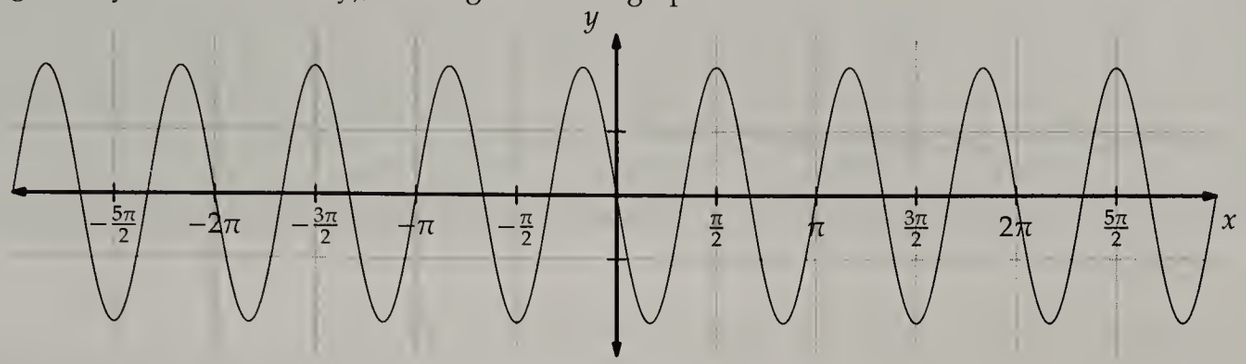


Figure 2.19: Graph $y = -2\sin 3x$

Finally, as explained in part (a), the graph of $y = -2\sin\left(3x - \frac{\pi}{4}\right)$ is a $\frac{\pi}{12}$ rightward shift of the graph of $y = -2\sin 3x$:

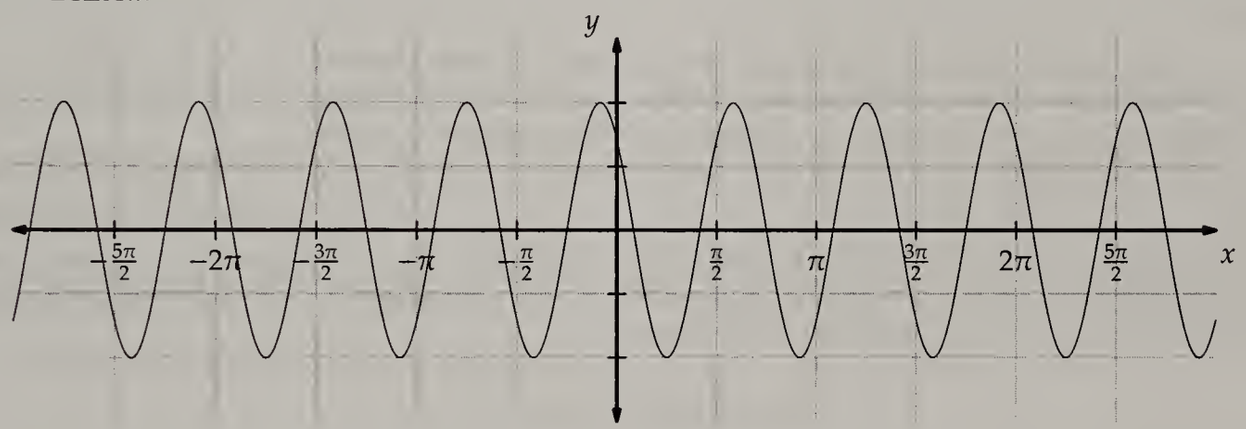


Figure 2.20: Graph $y = -2\sin\left(3x - \frac{\pi}{4}\right)$

□

Sidenote: One characteristic of trig functions that is used in physics is **frequency**. We often use frequency in conjunction with trigonometric functions of time. For example, we might have the function $A(t) = 3\sin(12\pi t)$, where t is in seconds. The frequency of this function is the number of periods the function completes in 1 second—that is, the frequency tells us how frequently the function repeats. See if you can figure out why this definition is equivalent to defining frequency as the reciprocal of the period.

One of the reasons we study complicated-looking trigonometric functions like $f(x) = 3\sin(4x + \pi)$ is that many processes in nature can be modeled with functions like these. You'll see such functions throughout your study of physics. For example, when you hear someone talk about frequency or wavelength in conjunction with sound, light, or even rays from outer space, they're describing a characteristic of a trigonometric function that can be used to model the sound, the light, or the space ray.

We say that $f(x)$ **varies sinusoidally** with a variable x, and that f is **sinusoidal**, if we can express f as $f(x) = a\sin(bx + c) + d$ for some constants a, b, c, and d.

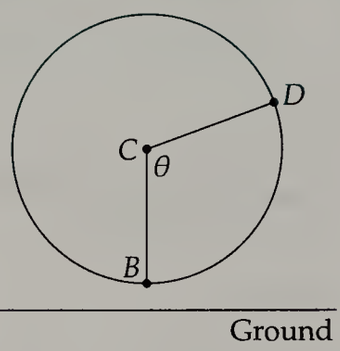
Problem 2.27: A Ferris wheel is a ride that consists of seats connected to the circumference of a large vertical wheel. As the wheel rotates, the riders are moved from the bottom of the wheel to the top of the wheel, and then back to the bottom. When you first get on a particular Ferris wheel, you are at the bottom of the ride, 3 feet off the ground. Once the Ferris wheel starts, it rotates around its center at a constant rate. Suppose that it takes you 8 seconds to reach the highest point of the ride, 63 feet off the ground.

(a) You start at the bottom of the ride. How high off the ground will you be 5 seconds after the ride starts? (Answer to the nearest tenth of a foot.)

(b) Find a function f such that $f(t)$ equals how high in feet you will be off the ground t seconds later.

Solution for Problem 2.27:

(a) In the diagram at right, you start at point B, and point C is the center of the Ferris wheel. After the Ferris wheel rotates an angle θ, you are at point D, as shown. Pretend for a moment that the Ferris wheel is the unit circle. Point B is directly below point C, so its coordinates would be $\left(\cos\frac{3\pi}{2}, \sin\frac{3\pi}{2}\right)$. Since point D is θ counterclockwise from B, its coordinates would be $\left(\cos\left(\frac{3\pi}{2}+\theta\right), \sin\left(\frac{3\pi}{2}+\theta\right)\right)$.



Next, we note that the Ferris wheel has radius 30 feet, since the distance from the bottom to the top of the wheel is 60 feet. So, $CB = 30$. Therefore, if we think of the Ferris wheel as a circle centered at the origin with radius 30, rather than radius 1, then the coordinates of B are $\left(30\cos\frac{3\pi}{2}, 30\sin\frac{3\pi}{2}\right)$ and the coordinates of D are $\left(30\cos\left(\frac{3\pi}{2}+\theta\right), 30\sin\left(\frac{3\pi}{2}+\theta\right)\right)$. The y-coordinate of D tells us how far you are above (when it is positive) or below (when it is negative) the center of the Ferris wheel after the wheel has turned an angle θ. Therefore, since the center of the Ferris wheel is 33 feet off the ground, your distance in feet from the ground after the wheel has turned an angle θ is given by the function

$$g(\theta) = 30\sin\left(\frac{3\pi}{2}+\theta\right) + 33.$$

So, we just have to figure out the angle θ after the wheel has turned for 5 seconds. The wheel rotates an angle π in 8 seconds, so in 5 seconds, it rotates by $5\pi/8$. Therefore, your distance from the ground after 5 seconds is

$$g\left(\frac{5\pi}{8}\right) = 30\sin\left(\frac{3\pi}{2}+\frac{5\pi}{8}\right) + 33 = 30\sin\frac{17\pi}{8} + 33 \approx 44.5 \text{ feet.}$$

(b) We convert our function in terms of the angle θ into a function in terms of the time t by finding the relationship between θ and t. Since the wheel turns an angle of π in 8 seconds, in t seconds it turns an angle $\theta = \frac{t}{8}\pi$. Therefore, the desired function is

$$f(t) = 30\sin\left(\frac{t\pi}{8}+\frac{3\pi}{2}\right) + 33.$$

This function looks a lot like the others we have studied in this section. We can confirm our answer by using our understanding of functions of the form $f(t) = a\sin(bt+c)+d$.

First, we note that the $a\sin(bt+c)$ term gives the vertical displacement between you and the center of the Ferris wheel ($a\sin(bt+c)$ is positive if you are above the center, and negative if you are below it). Your height from the ground is the sum of this displacement and the distance from the center of Ferris wheel to the ground, which is 33 feet. Therefore, we must have $d = 33$.

Next, because you go from 30 feet below the center to 30 feet above the center, we know that $a\sin(bt+c)$ varies from -30 to 30. That is, the amplitude of $a\sin(bt+c)$ is 30, so we can let $a = 30$.

The Ferris wheel makes a full rotation in 16 seconds, so the period of $a\sin(bt+c)$ must be 16 (assuming t is in seconds). Therefore, b must be $(2\pi)/16$, which equals $\pi/8$. Finally, we must account for the phase shift. You are at the bottom of the Ferris wheel at time $t = 0$, so $a\sin(bt+c)$ must be as small as possible when $t = 0$. This means we can let $c = 3\pi/2$, since $\sin(3\pi/2)$ gives us the smallest possible value of sine. Putting all these observations together gives us

$$f(t) = 30\sin\left(\frac{t\pi}{8} + \frac{3\pi}{2}\right) + 33,$$

which matches the function found earlier.

□

Exercises

2.4.1 Show that the amplitude of $f(x) = a\sin x$ is $|a|$.

2.4.2 Graph $y = 3 + \sin x$.

2.4.3

(a) Graph $y = \cos 2x$.

(b) Graph $y = -3\cos 2x$.

(c)★ Graph $y = -3\cos\left(2x - \frac{\pi}{3}\right)$.

2.4.4 Ted chooses two numbers t and u, and then graphs the function $g(x) = t\sin ux$. The resulting graph is shown below. If t is positive, find t and u.

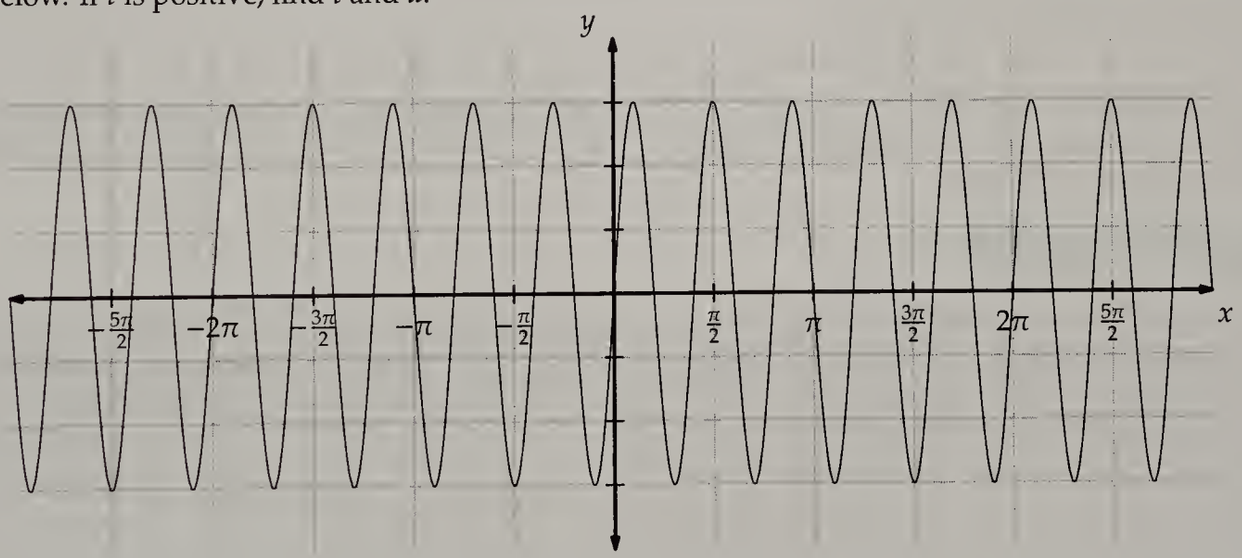


Figure 2.21: Graph for Exercise 2.4.4

2.4.5 Find the period, amplitude, and phase shift of $f(x) = -3\cos(2x + \pi) + 4$.

2.4.6 A block is attached to a long spring that is connected to the ceiling. The block is pulled away from the ceiling and then released, so that it bounces up and down, as the spring stretches and contracts, in a way that makes the distance from the block to the ceiling vary sinusoidally with time. Suppose the distance between the block and the ceiling varies between 1.2 meters and 1.8 meters, and that it takes 1 second for the block to go from high point to low point. How far from the ceiling is the block 2.7 seconds after it is at a high point?

2.4.7 Benny graphs a function of the form

$$f(x) = \cos(ax)$$

for some positive number a. Joon graphs a function of the form

$$f(x) = \sin(bx + c)$$

for some positive number b and real number c. They produce the same graph, which is shown below.

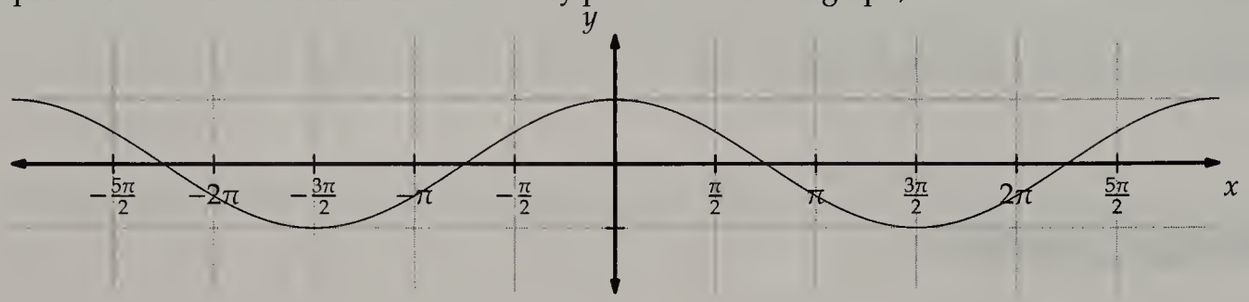


Figure 2.22: Graph for Exercise 2.4.7

(a) What is a?

(b) What is b?

(c) Find three possible values of c.

2.4.8 Is the function $f(x) = 2\sin x - \tan x$ periodic? If so, what is the period?

2.5 Inverse Trig Functions

So far in this chapter, we've mostly concerned ourselves with finding the output of trigonometric functions given certain inputs. In this section, we turn this around and try to find the inputs we must give to trig functions to produce certain desired outputs.

Problems

Problem 2.28: A cart is attached to a spring that is connected to a wall. The cart is pulled 8 feet away from the wall and then released, so that it rolls back and forth towards and away from the wall as the spring stretches and contracts. The distance from the cart to the wall varies sinusoidally with time. Suppose the distance between the cart and the wall varies between 4 feet and 8 feet, and that it takes 6 seconds for the cart to go from the point closest to the wall to the point that is farthest from the wall.

(a) Find a function that describes how far the cart is from the wall t seconds after it is released.

(b) How many seconds pass before the cart is exactly 5 feet from the wall for the first time?

(c) Find all values of t such that the cart is 5 feet from the wall after t seconds.

Problem 2.29:

(a) Explain why the function $f(x) = \cos x$ does not have an inverse.

(b) Suppose we define a function $g(x) = \cos x$ and restrict the domain of g to an interval $[a, b]$, where a and b are constants, such that g does have an inverse. What are valid choices for a and b? What are *good choices* for a and b?

Problem 2.28: A cart is attached to a spring that is connected to a wall. The cart is pulled 8 feet away from the wall and then released, so that it rolls back and forth towards and away from the wall as the spring stretches and contracts. The distance from the cart to the wall varies sinusoidally with time. Suppose the distance between the cart and the wall varies between 4 feet and 8 feet, and that it takes 6 seconds for the cart to go from the point closest to the wall to the point that is farthest from the wall.

(a) Find a function that describes how far the cart is from the wall t seconds after it is released.

(b) How many seconds pass before the cart is exactly 5 feet from the wall for the first time?

(c) Find all values of t such that the cart is 5 feet from the wall after t seconds.

Solution for Problem 2.28:

(a) Because the distance between the cart and the wall varies sinusoidally with time, this distance is

$$f(t) = a\sin(bt + c) + d,$$

where a, b, c, and d are constants. The middle point of the cart's path is 6 feet from the wall, and it varies 2 feet from this point in both directions. So, the amplitude of $a\sin(bt + c)$ is 2, and $f(t)$ must equal 6 when $a\sin(bt + c)$ is 0 (when the cart is at the midpoint of its path). Therefore, the function is $f(t) = 2\sin(bt + c) + 6$ for some values of b and c.

The cart starts 8 feet away from the wall, so we must have $f(0) = 8$. Therefore, we must have $\sin(bt+c) = 1$ when $t = 0$, which means $\sin c = 1$. This means we can let $c = \frac{\pi}{2}$, and we have

$$f(t) = 2\sin\left(bt + \frac{\pi}{2}\right) + 6.$$

(Note that we could let $c = \frac{5\pi}{2}$, or $-\frac{3\pi}{2}$, or any number that is a multiple of 2π from $\frac{\pi}{2}$.)

Finally, we turn to the last piece of information we have: the cart goes from the farthest point from the wall to the closest point in 6 seconds. This means that it goes from the farthest point to the closest and back to the farthest again in 12 seconds, so its period is 12 seconds. Since the period of $f(t)$ is $\frac{2\pi}{b}$ when b is positive, we must have $b = \pi/6$, and we have

$$f(t) = 2\sin\left(\frac{\pi}{6}t - \frac{\pi}{2}\right) + 4.$$

(b) We seek the smallest time t such that we have $f(t) = 5$, so we have

$$2\sin\left(\frac{\pi}{6}t - \frac{\pi}{2}\right) + 4 = 5.$$

Isolating the sine function, we have

$$\sin\left(\frac{\pi}{6}t - \frac{\pi}{2}\right) = \frac{1}{2}.$$

We now need to find an angle whose sine equals to $\frac{1}{2}$. Fortunately, we've seen $\sin\theta = \frac{1}{2}$ enough times by now to remember that $\sin\frac{\pi}{6} = \frac{1}{2}$. Therefore, if $\frac{\pi}{6}t - \frac{\pi}{2} = \frac{\pi}{6}$, then the equation above is satisfied. Solving for t gives $t = 4$. We still have to show that no smaller t is possible. The cart goes from 8 feet from the wall to 4 feet from the wall during the first 6 seconds, and the distance between the cart and the wall is strictly decreasing during that time. This means the cart can only be 5 feet from the wall once in those first 6 seconds. Since the cart is 5 feet from the wall when $t = 4$, this is smallest possible positive value of t.

(c) What's wrong with this solution:

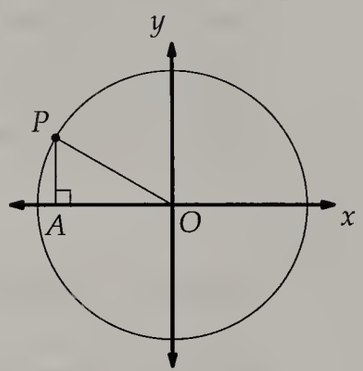
Bogus Solution: The period of sine is 2π, so the value of $\sin\left(\frac{\pi}{6}t - \frac{\pi}{2}\right)$ repeats every time the argument increases by 2π. Since we have $\sin\left(\frac{\pi}{6}t - \frac{\pi}{2}\right) = \frac{1}{2}$ when $\frac{\pi}{6}t - \frac{\pi}{2} = \frac{\pi}{6}$, we must have $\frac{\pi}{6}t - \frac{\pi}{2} = \frac{\pi}{6} + k(2\pi)$ for some nonnegative integer k. (No going backwards in time! We must have $k \geq 0$.) Therefore, the times at which the cart is 5 feet from the wall are $t = 4 + 12k$, for any nonnegative integer k.

We see the error in this reasoning by stepping back from the math and thinking about the problem.

WARNING!! When using math to model a problem, don't get so wrapped up in the math that you overlook what the problem is about. Keeping the original problem in mind can help you catch errors in your thinking.

The cart is 5 feet from the wall twice during each period—once as it approaches the wall and once as it moves away from the wall. The Bogus Solution only includes the times that are one full period from the first time the cart is 5 feet from the wall, which are the times during which the cart is approaching the wall. (We also could have caught this mistake by thinking about the unit circle—there are two points on the unit circle with y-coordinate equal to $\frac{1}{2}$, so there are two angles θ in each period of sine such that $\sin\theta = \frac{1}{2}$.)

Here are a couple ways to finish the problem:



Solution 1: Use Trigonometry. We can use the unit circle to find the other set of angles whose sine is $\frac{1}{2}$. In the diagram, point P is the second quadrant point on the unit circle with y-coordinate $\frac{1}{2}$. We draw altitude $\overline{PA}$ as shown; the length of this altitude equals the y-coordinate of P, so $PA = \frac{1}{2}$. Since leg $\overline{PA}$ of right triangle PAO has length $\frac{1}{2}$ the hypotenuse of PAO, the triangle is a 30-60-90 right triangle with $\angle POA = 30^\circ = \frac{\pi}{6}$. Therefore, P is the terminal point of $\pi - \frac{\pi}{6} = \frac{5\pi}{6}$.

So, in addition to the values of t that we found in the Bogus Solution, we also have $\sin\left(\frac{\pi}{6}t - \frac{\pi}{2}\right) = \frac{1}{2}$ when $\frac{\pi}{6}t - \frac{\pi}{2} = \frac{5\pi}{6} + (2\pi)k$, for any nonnegative integer k. Solving for t gives $t = 8 + 12k$, where k is a nonnegative integer. Combining this with the earlier values we found, the cart is 5 feet from the wall when $t = 4 + 12k$ seconds or $t = 8 + 12k$ seconds, for any nonnegative integer k.

Solution 2: Use Symmetry. Once again, we help ourselves by thinking about the original problem rather than staying focused on our equations. The cart goes from 8 feet away to 5 feet away in 4 seconds. Since it takes 6 seconds to go from 8 feet away to 4 feet away, it will take $6 - 4 = 2$ more seconds for the cart to complete its first trip from farthest point to nearest point. By symmetry, it will then take 2 more seconds to return to the point 5 feet from the wall, so $4 + 2 + 2 = 8$ seconds after the cart is released, it will again be 5 feet from the wall.

Since the period of the cart is 12 seconds and the two times in the first period at which the cart is 5 feet from the wall are at $t = 4$ seconds and $t = 8$ seconds, the values of t for which the cart are 5 feet from the wall are $t = 4 + 12k$ seconds or $t = 8 + 12k$ seconds, for any nonnegative integer k.

□

Problem 2.29:

(a) Explain why the function $f(x) = \cos x$ does not have an inverse.

(b) Suppose we define a function $g(x) = \cos x$ and restrict the domain of g to an interval $[a, b]$, where a and b are constants, such that g does have an inverse. What are valid choices for a and b? What are *good choices* for a and b?

Solution for Problem 2.29:

(a) A function is invertible if and only if there is exactly one input that produces each possible output from the function. Since $\cos 0 = 1$ and $\cos 2\pi = 1$, the function $f(x) = \cos x$ has two inputs that produce the same output, so it is not invertible. Moreover, since $\cos x$ is periodic, all possible outputs of $f(x)$ occur for infinitely many inputs.

(b) Having noted that $\cos x$ is periodic, we might think that if we restrict the domain of g to a single period of $\cos x$, then we will be able to invert it. But this fails, too, because there are many values of y such that $y = \cos x$ for two different values of x in the same period. For example, if we consider the period consisting of the interval $[0, 2\pi)$, we have $\cos \frac{\pi}{2} = \cos \frac{3\pi}{2}$ and $\cos \frac{\pi}{4} = \cos \frac{7\pi}{4}$, and so on. So, we can't take a full period as the domain of g.

If we take a look at the graph of $y = \cos x$, we can quickly see what sorts of intervals we can take as the domain of g such that g is invertible.

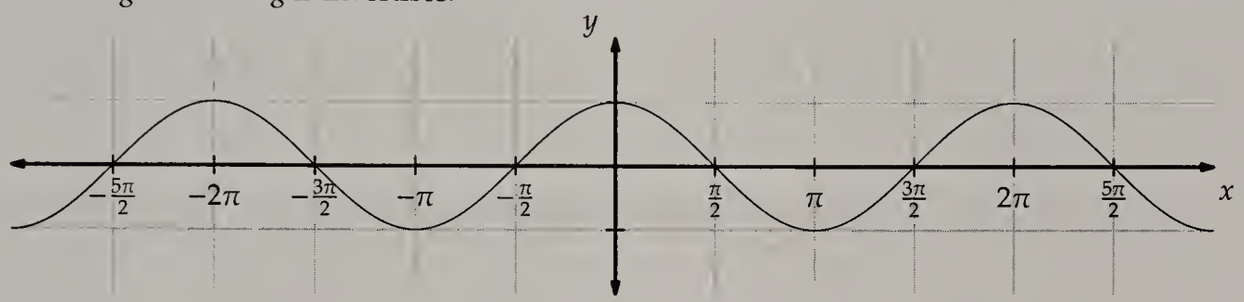


Figure 2.23: Graph of $y = \cos x$

If a function is invertible, then no two points on its graph can have the same y-coordinate. So, we can choose a suitable domain such that g is invertible by picking an interval such that the graph of $y = \cos x$ has no two points with the same y-coordinate when x is in the interval.

Now, we have some choices to make—there are infinitely many ways we can choose this interval. We could choose $\left[0, \frac{\pi}{2}\right]$, for example. But this wouldn't be such a useful function. We couldn't use it for cases in which cosine is negative, for example. We'd like to choose an interval that provides all the possible outputs of $\cos x$. In other words, we'd like the range of g to be $[-1, 1]$, because then it is the same as the range of $\cos x$ when there are no restrictions on the domain. Looking at the graph of $y = \cos x$, or thinking about the unit circle, we see that the interval $[0, \pi]$ fits the bill nicely, since $\cos x$ takes on every possible value of $\cos x$ once and exactly once for $x \in [0, \pi]$.

The interval $[0, \pi]$ isn't the only possible domain for g such that g is invertible and has range $[-1, 1]$. We could have chosen $[\pi, 2\pi]$, or even a union of intervals such as $\left[0, \frac{\pi}{2}\right] \cup \left[3\pi, \frac{7\pi}{2}\right)$, where $g^{-1}(x)$ is in the interval $\left[0, \frac{\pi}{2}\right]$ if x is nonnegative and $g^{-1}(x)$ is in $\left[3\pi, \frac{7\pi}{2}\right)$ if x is negative.

Any of these choices is perfectly fine for the domain of g. By convention, when we wish to have a version of cosine that is invertible, we restrict the domain to $[0, \pi]$. As we've seen, this allows us to find a unique input for any possible output of cosine. Moreover, $[0, \pi]$ is a single interval rather than a disjoint combination of intervals, and it's the easiest single interval to remember, perhaps because it's "natural" to start from 0 and go up.

□

We have a special family of functions that return an angle given the value of a trigonometric function of that angle. These are the **inverse trigonometric functions**, and they are the inverses of the trigonometric functions when the domains of these functions are suitably restricted to make the resulting function invertible.

For example, we define the function $f(x) = \arccos x$ to output the angle in the interval $[0, \pi]$ whose cosine is x. The domain of $\arccos x$ is $[-1, 1]$, since the cosine of an angle must be in this interval. As described in Problem 2.29,

$[0, \pi]$ isn't the only interval we could have chosen to be the range of $\arccos x$, but it is the convention we typically use for the reasons explained in Problem 2.29. So, while there are infinitely many angles x such that $\cos x = \frac{1}{2}$, the value of $\arccos \frac{1}{2}$ is the angle in the interval $[0, \pi]$ such that $\cos x = \frac{1}{2}$. Therefore, we have $\arccos \frac{1}{2} = \frac{\pi}{3}$.

With $\arccos x$ defined in this manner, we have $\cos(\arccos x) = x$ for all x in the domain of $\arccos x$, and $\arccos(\cos x) = x$ for $0 \leq x \leq \pi$. Therefore, we can think of $\arccos x$ as the inverse function of $\cos x$ in that the functions "undo" each other, where we restrict the domain of $\cos x$ to $[0, \pi]$ in order to define $\arccos x$.

We call the function $\arccos x$ "**arccosine**" or "**inverse cosine.**" Sometimes this function is denoted $\cos^{-1} x$. We'll use both $\arccos x$ and $\cos^{-1} x$ interchangeably in this section so you become used to both notations, but for the rest of the text we'll prefer $\arccos x$ to avoid confusion, since some might interpret $\cos^{-1} x$ as the reciprocal of $\cos x$. (We have the function $\sec x$, which equals the reciprocal of $\cos x$, to help us avoid this confusion.)

There's nothing special about cosine; we have inverse functions for each of the primary trigonometric functions. Just as with $\arccos x$, we must place special restrictions on the ranges of the inverses of the other trig functions. Also just as with $\arccos x$, the specific choices we make for these restrictions are matters of convention, and not the only choices we could have made. Below is a table of the inverse trig functions; as you read over it, think about why we choose the indicated ranges.

Trig function	Inverse Trig Function	Domain of Inverse	Range of Inverse
$\cos x$	$\arccos x$	$[-1, 1]$	$[0, \pi]$
$\sin x$	$\arcsin x$	$[-1, 1]$	$\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$
$\tan x$	$\arctan x$	$\mathbb{R}$	$\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$
$\cot x$	$\operatorname{arccot} x$	$\mathbb{R}$	$(0, \pi)$
$\sec x$	$\operatorname{arcsec} x$	$(-\infty, -1] \cup [1, +\infty)$	$\left[0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right]$
$\csc x$	$\operatorname{arccsc} x$	$(-\infty, -1] \cup [1, +\infty)$	$\left[-\frac{\pi}{2}, 0\right) \cup \left(0, \frac{\pi}{2}\right]$

Don't memorize all these domains and ranges right now. While working on the following problems, you'll learn more about why we must restrict the domains and ranges as described in the table above. Most of the restrictions will become obvious to you, rather than something to be memorized.

Problems

Problem 2.30: Evaluate each of the following:

(a) $\arcsin 0$ (c) $\arctan 1$ (e) $\operatorname{arccot} \sqrt{3}$

(b) $\cos^{-1} \frac{\sqrt{3}}{2}$ (d) $\arcsin\left(-\frac{\sqrt{2}}{2}\right)$ (f) $\sec^{-1} \frac{2\sqrt{3}}{3}$

Problem 2.31: If x is nonzero and $\arccos x = \operatorname{arcsec} y$, then find y in terms of x.

Problem 2.32:

(a) Why is $[-1, 1]$ the domain of $\arccos x$, but the domain of $\arctan x$ is $\mathbb{R}$?

(b) We use $[0, \pi]$ as the range of $\arccos x$. Why can't the range be a smaller or larger interval?

(c) Why can't we also use $[0, \pi]$ as the range of $\arcsin x$?

(d) Why must we exclude the endpoints of the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ from the range of $\arctan x$?

Problem 2.33: Evaluate each of the following without a calculator:

(a) $\arcsin(\sin(39\pi/5))$ (b) $\arcsin(\sin(37\pi/5))$

Problem 2.34:

(a) Evaluate $\sin(\arccos 0.6)$ without a calculator.

(b) Evaluate $\sin(\arccos(-0.6))$ without a calculator.

Problem 2.35: Let $y = \arcsin x$. In this problem, we find $\tan y$ in terms of x.

(a) Write the equation $y = \arcsin x$ as an equivalent equation without an inverse trig function.

(b) Write $\tan y$ in terms of sine and cosine.

(c) Find $\cos y$ in terms of x.

(d) Find $\tan y$ in terms of x.

Problem 2.36: Graph $y = \arcsin x$. How is this graph related to the graph of $y = \sin x$?

Problem 2.37: In this problem, we find all t such that $2\cos\left(3t - \frac{\pi}{4}\right) = \sqrt{3}$.

(a) Find all t such that $\cos t = \frac{\sqrt{3}}{2}$.

(b) Find all t such that $\cos\left(t - \frac{\pi}{4}\right) = \frac{\sqrt{3}}{2}$.

(c) Find all t such that $2\cos\left(3t - \frac{\pi}{4}\right) = \sqrt{3}$.

Problem 2.38: What angles ϕ satisfy $\sec\phi\tan\phi = -\frac{2}{3}$?

Problem 2.30: Evaluate each of the following:

(a) $\arcsin 0$ (c) $\arctan 1$ (e) $\operatorname{arccot}\sqrt{3}$

(b) $\cos^{-1}\frac{\sqrt{3}}{2}$ (d) $\arcsin\left(-\frac{\sqrt{2}}{2}\right)$ (f) $\sec^{-1}\frac{2\sqrt{3}}{3}$

Solution for Problem 2.30:

(a) Since $\sin 0 = 0$, we have $\arcsin 0 = 0$.

(b) Since $\cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}$, we have $\cos^{-1}\frac{\sqrt{3}}{2} = \frac{\pi}{6}$.

(c) If $x = \arctan 1$, then $\sin x$ and $\cos x$ must be equal, so that $\tan x$ will equal 1. The only angle in the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ for which this is the case is $\frac{\pi}{4}$, so $\arctan 1 = \frac{\pi}{4}$.

(d) We know that $\sin x$ equals $\frac{\sqrt{2}}{2}$ or $-\frac{\sqrt{2}}{2}$ for odd integer multiples of $\frac{\pi}{4}$. Here, we want the one in the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ for which sine is negative, so we have $\arcsin\left(-\frac{\sqrt{2}}{2}\right) = -\frac{\pi}{4}$.

(e) If $\cot x = \sqrt{3}$, then we have $\frac{\cos x}{\sin x} = \sqrt{3}$. The presence of $\sqrt{3}$ suggests testing $x = \frac{\pi}{6}$ and $x = \frac{\pi}{3}$. The former works, so $\operatorname{arccot}\sqrt{3} = \frac{\pi}{6}$.

If we hadn't thought of testing $x = \frac{\pi}{6}$, all is not lost. From $\frac{\cos x}{\sin x} = \sqrt{3}$, we have $\cos x = \sqrt{3}\sin x$. The identity $\sin^2 x + \cos^2 x = 1$ also relates $\sin x$ and $\cos x$. Substituting $\cos x = \sqrt{3}\sin x$ into this gives $4\sin^2 x = 1$, so $\sin^2 x = \frac{1}{4}$. Therefore, we have $\sin x = \pm\frac{1}{2}$. Since the range of $\operatorname{arccot} t$ is $(0, \pi)$ and $\cot x$ is negative for

$\frac{\pi}{2} < x < \pi$, we must have $0 < \text{arccot}\,\sqrt{3} < \frac{\pi}{2}$. This means $\sin x > 0$, so $\sin x = \frac{1}{2}$ and we have $x = \frac{\pi}{6}$, as before.

(f) We seek the angle x such that $\sec x = \frac{2\sqrt{3}}{3}$. We're a lot more comfortable with sine and cosine than we are with secant, so we rewrite this equation as

$$\frac{1}{\cos x} = \frac{2\sqrt{3}}{3}.$$

Taking the reciprocal of both sides gives us

$$\cos x = \frac{3}{2\sqrt{3}} = \frac{\sqrt{3}}{2},$$

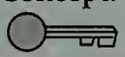
and we're in familiar territory. In fact, this is now the same problem as part (b), so we have $\text{arcsec}\,\frac{2\sqrt{3}}{3} = \frac{\pi}{6}$.

□

Problem 2.31: If x is nonzero and $\arccos x = \text{arcsec}\, y$, then find y in terms of x.

Solution for Problem 2.31: Let $\theta = \arccos x = \text{arcsec}\, y$, so we have $\cos\theta = x$ and $\sec\theta = y$. Since $x \neq 0$, we can take the reciprocal of both sides of $x = \cos\theta$ to get $\frac{1}{x} = \frac{1}{\cos\theta} = \sec\theta = y$. Therefore, we have $y = \frac{1}{x}$. □

A key step in our solution to Problem 2.31 is assigning a variable to the angle that equals the inverse trig functions in the problem, and then rewriting the information in the problem in terms of trig functions instead of inverse trig functions.

Concept:  Some problems involving inverse trig functions can be most easily solved by rewriting the problem in terms of the corresponding trig functions.

Now that we have a little practice working with inverse trig functions, let's take a look at why we restrict their domains and ranges the way we do.

Problem 2.32:

(a) Why is $[-1, 1]$ the domain of $\arccos x$, but the domain of $\arctan x$ is $\mathbb{R}$?

(b) We use $[0, \pi]$ as the range of $\arccos x$. Why can't the range be a smaller or larger interval?

(c) Why can't we also use $[0, \pi]$ as the range of $\arcsin x$?

(d) Why must we exclude the endpoints of the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ from the range of $\arctan x$?

Solution for Problem 2.32:

(a) The value of $\arccos x$ is the angle whose cosine is x. Since cosine cannot be larger than 1 or less than -1, we must restrict x in $\arccos x$ to be in the interval $[-1, 1]$. Meanwhile, for any real number x, there's some angle whose tangent is x. Therefore, the domain of $\arctan x$ is $\mathbb{R}$.

(b) Any possible range of $\arccos x$ must include angles with terminal points $(1, 0)$ and $(-1, 0)$ on the unit circle. These two points are terminal points of angles that are π apart. Moreover, as an angle increases from 0 to π, the x-coordinate of its terminal point decreases from 1 to -1 and is never the same for two different angles. To include the angles 0 and π, and all the angles in between, we need an interval of at least length π. To see why the range of $\arccos x$ cannot be an interval longer than π, note that we cannot allow two angles with the same cosine in the range. Since the circumference of the unit circle covers a range of 2π, any interval greater than π will cover more than half the unit circle. Such an interval will therefore include two points with the same x-coordinate, and these two points correspond to angles with the same cosine.

(c) Using $[0, \pi]$ as the range creates two problems. First, what would $\arcsin 0.5$ be? Both $\sin \frac{\pi}{6}$ and $\sin \frac{5\pi}{6}$ equal 0.5. Second, what would $\arcsin(-1)$ be? There are no angles in the interval $[0, \pi]$ that have a negative sine.

(d) Tangent is not defined for angles of $\frac{\pi}{2}$ or $-\frac{\pi}{2}$, so we must exclude these values from the range of $\arctan x$, since there is no value of x such that $\tan \frac{\pi}{2} = x$ or $\tan\left(-\frac{\pi}{2}\right) = x$.

□

Problem 2.33: Evaluate each of the following without a calculator:

(a) $\arcsin(\sin(39\pi/5))$ (b) $\arcsin(\sin(37\pi/5))$

Solution for Problem 2.33:

(a) What's wrong with this answer:

Bogus Solution: Since $\arcsin(\sin x) = x$, we have $\arcsin(\sin(39\pi/5)) = 39\pi/5$.

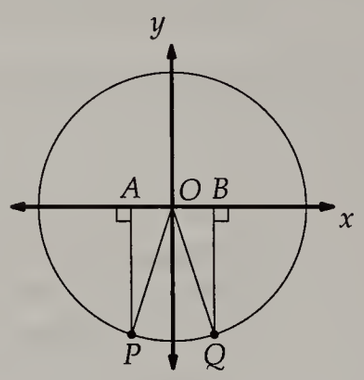
The mistake here is that the range of $\arcsin x$ is $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, and $\frac{39\pi}{5}$ is not in this interval. We only have $\arcsin(\sin x) = x$ if $x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. So, we have to find the angle in the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ whose sine equals $\sin \frac{39\pi}{5}$.

While we haven't learned yet how to evaluate $\sin \frac{39\pi}{5}$ without a calculator, we do know how to find other angles that have the same sine as the sine of $\frac{39\pi}{5}$. Subtracting multiples of 2π from $\frac{39\pi}{5}$ produces other angles whose sine equals $\sin \frac{39\pi}{5}$, since subtracting 2π from an angle does not change its sine. We find that $\frac{39\pi}{5} - 4 \cdot (2\pi) = -\frac{\pi}{5}$, so $\sin \frac{39\pi}{5} = \sin\left(-\frac{\pi}{5}\right)$. But we don't know how to evaluate $\sin\left(-\frac{\pi}{5}\right)$, either. However, $-\frac{\pi}{5}$ is in the range of $\arcsin x$, so whatever angle $\sin\left(-\frac{\pi}{5}\right)$ is, when we take the arcsine of that angle, we'll get $-\frac{\pi}{5}$ back:

$$\arcsin\left(\sin \frac{39\pi}{5}\right) = \arcsin\left(\sin\left(-\frac{\pi}{5}\right)\right) = -\frac{\pi}{5}.$$

(b) If we try the same tactic as in the previous part, we find that $\frac{37\pi}{5} - 3(2\pi) = \frac{7\pi}{5}$ and $\frac{37\pi}{5} - 4(2\pi) = -\frac{3\pi}{5}$. Uh-oh, neither $\frac{7\pi}{5}$ nor $-\frac{3\pi}{5}$ is in the range of $\arcsin x$. There isn't an angle in the range of $\arcsin x$ with the same terminal point as the angle $\frac{37\pi}{5}$. Does that mean there's no solution?

No! We don't need an angle in the range of $\arcsin x$ with the same terminal point as $\frac{7\pi}{5}$; we need an angle in this range with the same *sine*. At right we have the unit circle, and P is the terminal point of $\frac{37\pi}{5}$. There is another point, Q, on the unit circle with the same y-coordinate as P. This point is the terminal point of an angle whose sine equals $\sin \frac{37\pi}{5}$. Congruent right triangles $\triangle POA$ and $\triangle QOB$ give us $\angle QOB = \angle POA$. Since P is the terminal point of $-\frac{3\pi}{5}$, we have $\angle POB = \frac{3\pi}{5}$, so $\angle POA = \pi - \frac{3\pi}{5} = \frac{2\pi}{5}$, which means $\angle QOB = \frac{2\pi}{5}$. Therefore, Q is the terminal point of $-\frac{2\pi}{5}$, and we have



$$\arcsin\left(\sin \frac{37\pi}{5}\right) = \arcsin\left(\sin\left(-\frac{3\pi}{5}\right)\right) = \arcsin\left(\sin\left(-\frac{2\pi}{5}\right)\right) = -\frac{2\pi}{5}.$$

□

Problem 2.34:

(a) Evaluate $\sin(\arccos 0.6)$ without a calculator.

(b) Evaluate $\sin(\arccos(-0.6))$ without a calculator.

Solution for Problem 2.34:

(a) If we could find the angle θ with $0 \le \theta \le \pi$ such that $\cos\theta = 0.6$, then our problem would become finding $\sin\theta$. But without a calculator, it's not clear how to find the angle whose cosine is 0.6.

While we can't evaluate $\arccos 0.6$, we do at least have another way to view the problem. Instead of "Evaluate $\sin(\arccos 0.6)$," we can view the problem as

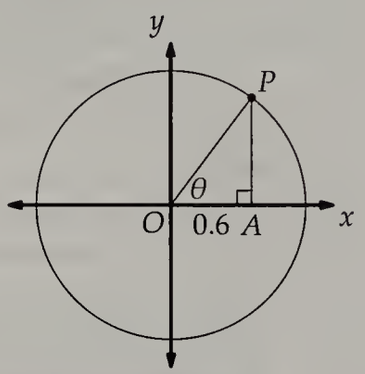
$$\text{We have } \cos\theta = 0.6, \text{ where } 0 \le \theta \le \pi. \text{ Find } \sin\theta.$$

This at least has the advantage of getting rid of the confusing inverse trig function, and leaves us with regular trig functions, with which we're more familiar.

Concept: Rewriting problems involving inverse trig functions as problems involving only regular trig functions can sometimes make the solution easier to find.

Our problem now is to find $\sin\theta$ given $\cos\theta$, so we look for a way to relate the two. Here are two ways to do so:

Method 1: The Unit Circle. Since $\theta = \arccos 0.6$, it is a first quadrant angle, as shown in the diagram at right. The Pythagorean Theorem then gives us $AO^2 + PA^2 = PO^2$, so $PA = \sqrt{PO^2 - AO^2} = \sqrt{1^2 - 0.6^2} = 0.8$. (We might also have recalled the Pythagorean triple $\{3,4,5\}$; dividing these by 5 gives $\{0.6, 0.8, 1\}$.) We then have $\sin\theta = PA = 0.8$.



Method 2: Use a Trig Identity. Our first approach is essentially equivalent to using the trig identity $\sin^2\theta + \cos^2\theta = 1$. Since $\cos\theta = 0.6$, we have $\sin^2\theta = 1 - (0.6)^2 = 0.64$.

We have to be careful to avoid this mistake:

Bogus Solution: Taking the square root of both sides gives $\sin\theta = \pm 0.8$, so there are two solutions.

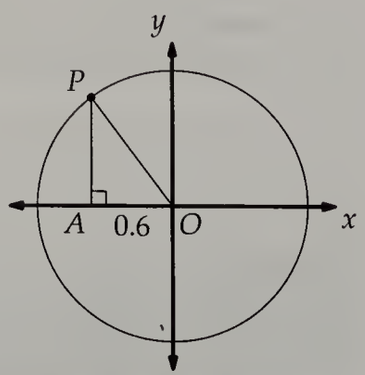
Looking back at the original problem, we know immediately that something is wrong here because $\sin(\arccos x)$ is a function. Therefore, for any value of x in the domain of this function, there is at exactly one value of $\sin(\arccos x)$. To determine which of 0.8 and -0.8 is the value of $\sin(\arccos 0.6)$, we must remember that $\theta = \arccos 0.6$. This means θ is restricted to the range of $\arccos x$, which is the interval $[0, \pi]$. Sine is nonnegative for all angles in this interval, so we discard the negative solution and have $\sin\theta = 0.8$.

(b) We follow our steps from part (a), and recast the problem as:

$$\text{We have } \cos\theta = -0.6, \text{ where } 0 \le \theta \le \pi. \text{ Find } \sin\theta.$$

As before, we can finish with either the unit circle or a trig identity.

Method 1: The Unit Circle. Since $\cos\theta = -0.6$ and θ is in the interval $[0, \pi]$, the angle θ is in the second quadrant, as shown. From $\cos\theta = -0.6$, we have $OA = 0.6$. Therefore, we have $PA = \sqrt{PO^2 - OA^2} = 0.8$, so $\sin\theta = 0.8$.



Method 2: Trig Identity. As in part (a), we use $\sin^2\theta + \cos^2\theta = 1$, from which we find $\sin^2\theta = 1 - \cos^2\theta = 1 - 0.36 = 0.64$. Since $0 \le \theta \le \pi$ and $\sin\theta$ is nonnegative for all angles θ in this range, $\sin^2\theta = 0.64$ gives us $\sin\theta = 0.8$.

□

Problem 2.35: Let $y = \arcsin x$. Find $\tan y$ in terms of x.

Solution for Problem 2.35: We start by getting rid of the inverse trig function. We write $y = \arcsin x$ as $\sin y = x$, where $y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Now, we must relate $\tan y$ to $\sin y$ in order to express $\tan y$ in terms of x. Unfortunately, we don't have a nice neat identity like we have to relate $\sin y$ and $\cos y$. But we can write $\tan y$ in terms of $\sin y$ and $\cos y$, and we already have $\sin y$ in terms of x:

$$\tan y = \frac{\sin y}{\cos y} = \frac{x}{\cos y}.$$

We now just have to express $\cos y$ in terms of x. We know how to relate $\cos y$ to $\sin y$. From $\sin^2 y + \cos^2 y = 1$, we have

$$\cos^2 y = 1 - \sin^2 y = 1 - x^2.$$

Once again, we face the puzzle of choosing the positive or negative root. Since $y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ (because y is in the range of $\arcsin x$) and $\cos y$ is nonnegative for all y in this interval, we have $\cos y = \sqrt{1 - x^2}$, so

$$\tan y = \frac{x}{\cos y} = \frac{x}{\sqrt{1 - x^2}}.$$

□

After rewriting the problem without inverse trig functions, our key problem solving strategy in this solution is:

Concept: When stuck on a trig problem involving tangent, cotangent, secant, and or cosecant, try rewriting the problem in terms of sine and cosine.

Problem 2.36: Graph $y = \arcsin x$. How is this graph related to the graph of $y = \sin x$?

Solution for Problem 2.36: The domain of $\arcsin x$ is $[-1, 1]$ and the range is $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, so the entire graph is contained in the region $-1 \le x \le 1$, $-\frac{\pi}{2} \le y \le \frac{\pi}{2}$. When $x = -1$, we have $\arcsin x = -\frac{\pi}{2}$. As x increases from -1 to 1, $\arcsin x$ increases from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$.

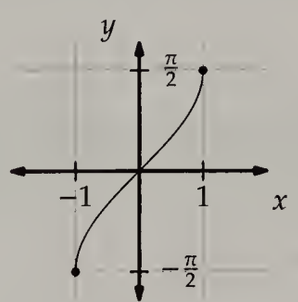


Figure 2.24: Graph of $y = \arcsin x$

The shape of this graph shouldn't be a surprise; on page 23, we explained that the graph of a function's inverse is the reflection of the function's graph over the graph of $y = x$. We see this relationship when we graph $y = \sin x$ for $-\frac{\pi}{2} \le x \le \frac{\pi}{2}$ and graph $y = \arcsin x$ on the same Cartesian plane:

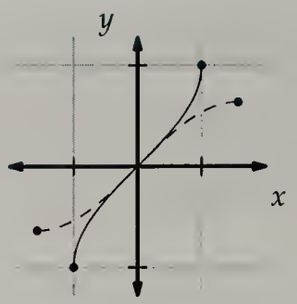


Figure 2.25: Graphs of $y = \arcsin x$ (solid) and a portion of $y = \sin x$ (dashed)

□

Now that we know how to handle inverse trigonometric functions, let's tackle a couple equations in which we must "undo" trig functions much like we do when we evaluate inverse trig functions.

Problem 2.37: Find all t such that $2\cos\left(3t - \frac{\pi}{4}\right) = \sqrt{3}$.

Solution for Problem 2.37: The cosine expression looks pretty complicated. Let's try to get a handle on the problem by solving a simpler version of the problem first.

Concept: Solving a simpler version of a complicated problem can help solve the original complex problem.

Since it's the complexity of the trig function that is the difficulty here, let's replace the left side with a simpler trig function, like $\cos x$. This gives us

$$\cos x = \sqrt{3}.$$

Uh-oh. That equation has no solutions. Did we change the problem too much, or does the original equation have no solutions?

Looking back at the original equation, $2\cos\left(3t - \frac{\pi}{4}\right) = \sqrt{3}$, it's the $\left(3t - \frac{\pi}{4}\right)$ that makes the cosine complicated, not the coefficient 2. We can divide both sides by 2 to isolate the cosine:

$$\cos\left(3t - \frac{\pi}{4}\right) = \frac{\sqrt{3}}{2}.$$

Concept: In equations involving trig functions, try to isolate trig functions by manipulating the equation so that you have a trig function equal to a constant.

Now, our "simpler version" of the problem makes more sense:

$$\cos x = \frac{\sqrt{3}}{2}.$$

As in Problem 2.29, there are infinitely many solutions to this equation. We have $\cos x = \frac{\sqrt{3}}{2}$ for $x = \frac{\pi}{6}$, for $x = \frac{11\pi}{6}$, and for any angle that is a multiple of 2π more or less than either of these angles. We can write these solutions as $x = \frac{\pi}{6} + 2k\pi$ or $x = \frac{11\pi}{6} + 2k\pi$, for any integer k.

But how does this help us with

$$\cos\left(3t - \frac{\pi}{4}\right) = \frac{\sqrt{3}}{2}?$$

The only difference between this and $\cos x = \frac{\sqrt{3}}{2}$ is that we've replaced x with $\left(3t - \frac{\pi}{4}\right)$. So, to solve our complicated problem, we replace x with $\left(3t - \frac{\pi}{4}\right)$ in our solution to the simpler problem. Since "$x = \frac{\pi}{6} + 2k\pi$ or $x = \frac{11\pi}{6} + 2k\pi$, for any integer k," are the solutions to $\cos x = \frac{\sqrt{3}}{2}$, the solutions to $\cos\left(3t - \frac{\pi}{4}\right) = \frac{\sqrt{3}}{2}$ are the values of t such that

$$3t - \frac{\pi}{4} = \frac{\pi}{6} + 2k\pi \qquad \text{or} \qquad 3t - \frac{\pi}{4} = \frac{11\pi}{6} + 2k\pi,$$

for any integer k. Isolating t in each of these gives us the solutions

$$t = \frac{5\pi}{36} + \frac{2k\pi}{3} \qquad \text{or} \qquad t = \frac{25\pi}{36} + \frac{2k\pi}{3},$$

for any integer k.

As an Exercise, you'll explore the fact that there are other forms in which we can express these solutions. □

Problem 2.38: What angles ϕ satisfy $\sec\phi \tan\phi = -\frac{2}{3}$?

Solution for Problem 2.38: This problem is more complicated than the previous problem for a couple reasons. First, it's not immediately possible to rewrite the equation as a trig function equal to a constant. Second, the trig functions are secant and tangent, which are less familiar than sine and cosine. We'll tackle the the second difficulty first, by rewriting the problem in terms of sine and cosine:

$$\frac{1}{\cos\phi} \cdot \frac{\sin\phi}{\cos\phi} = -\frac{2}{3}.$$

Multiplying both sides by $3\cos^2\phi$ to get rid of the fractions gives

$$3\sin\phi = -2\cos^2\phi.$$

Now what?

Concept: Trig equations with a single type of trig function are often easier to deal with than equations with two different types of trig functions.

We can write $3\sin\phi = -2\cos^2\phi$ in terms of just $\sin\phi$ by noting that $\cos^2\phi = 1 - \sin^2\phi$, so our equation is $3\sin\phi = -2(1 - \sin^2\phi)$. Rearranging this equation gives

$$2\sin^2\phi - 3\sin\phi - 2 = 0.$$

This is a quadratic equation in $\sin\phi$. (If you don't see why, let $y = \sin\phi$, which makes the equation $2y^2 - 3y - 2 = 0$.) Factoring this equation gives

$$(2\sin\phi + 1)(\sin\phi - 2) = 0.$$

Therefore, we must have $2\sin\phi + 1 = 0$, which gives $\sin\phi = -\frac{1}{2}$, or $\sin\phi - 2 = 0$, which gives $\sin\phi = 2$. We have succeeded in manipulating the equation into a single trig function set equal to a constant.

We can discard $\sin\phi = 2$, since sine cannot be more than 1. To solve $\sin\phi = -\frac{1}{2}$, we first find the two solutions with $0 \le \phi < 2\pi$, which correspond to the two points on the unit circle with y-coordinate $-\frac{1}{2}$. These are $\phi = \frac{7\pi}{6}$ and $\phi = \frac{11\pi}{6}$. Checking these confirms that both do satisfy the original equation. We can add any multiple of 2π to these angles, so our full solution is $\phi = \frac{7\pi}{6} + 2k\pi$ or $\phi = \frac{11\pi}{6} + 2k\pi$ for any integer k. □

Exercises

2.5.1 Evaluate each of the following:

(a) $\arcsin 0.5$ (c) $\tan^{-1}(-1)$ (e) $\operatorname{arccot} \frac{\sqrt{3}}{3}$

(b) $\arccos\left(-\frac{\sqrt{2}}{2}\right)$ (d) $\arccos 0$ (f) $\csc^{-1}(-2)$

2.5.2 Evaluate $\cos(\arcsin 0.3)$ without using a calculator.

2.5.3 Evaluate $\arctan\left(\tan \frac{63\pi}{5}\right)$.

2.5.4 Why do we choose $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ as the range of $\arctan x$ instead of using $[0, \pi)$ as the range?

2.5.5 Find all values of x for which $\sin(\arccos x) \leq 0$. Are there any values of x for which we have $\cos(\arcsin x) < 0$?

2.5.6 Let $y = \arctan x$. Find $\sin y$ in terms of x.

2.5.7 Kira solved Problem 2.37 and produced the solutions

$$t = \frac{5\pi}{36} + \frac{2k\pi}{3} \quad \text{or} \quad t = \frac{\pi}{36} + \frac{2k\pi}{3},$$

for any integer k. The second set of solutions appears to be subtly different than our solution to Problem 2.37. Did Kira make a mistake?

2.6 Summary

The **unit circle** is the circle with radius 1 centered at the origin of the Cartesian plane. The **terminal point** of an angle θ is the point that results when the point $(1, 0)$ is rotated counterclockwise by θ about the origin.

Definition: The **cosine** of an angle θ is the x-coordinate of the terminal point of θ. We denote the cosine of θ as $\cos\theta$.

The **sine** of an angle θ is the y-coordinate of the terminal point of θ. We denote the sine of θ as $\sin\theta$.

The **tangent** of an angle θ is the ratio of the y-coordinate to the x-coordinate of the terminal point of θ. We denote the tangent of θ as $\tan\theta$.

When $\cos\theta \neq 0$, we can express $\tan\theta$ as $\frac{\sin\theta}{\cos\theta}$. If $\cos\theta = 0$, then $\tan\theta$ is undefined.

The graphs of sine, cosine, and tangent are below.

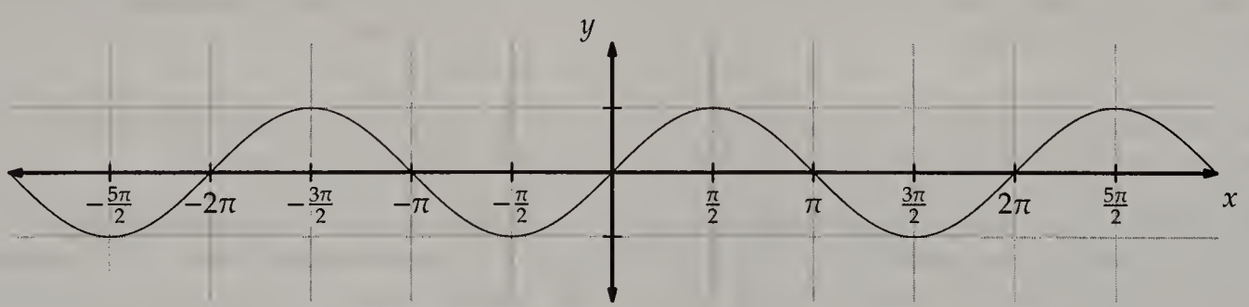


Figure 2.26: Graph of $y = \sin x$

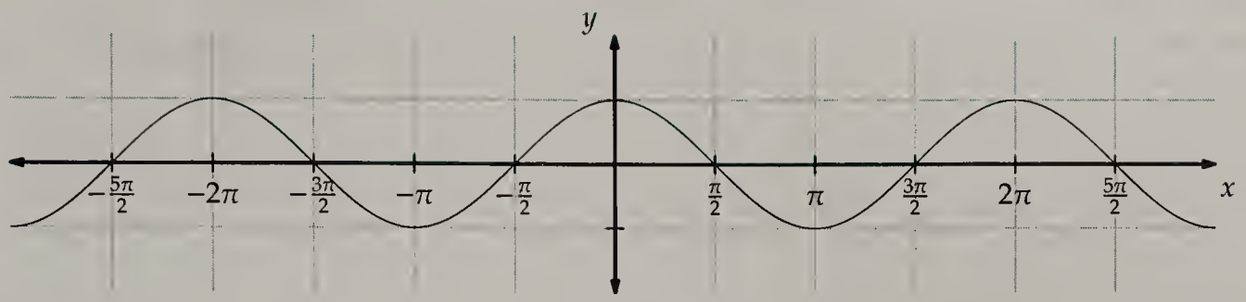


Figure 2.27: Graph of $y = \cos x$

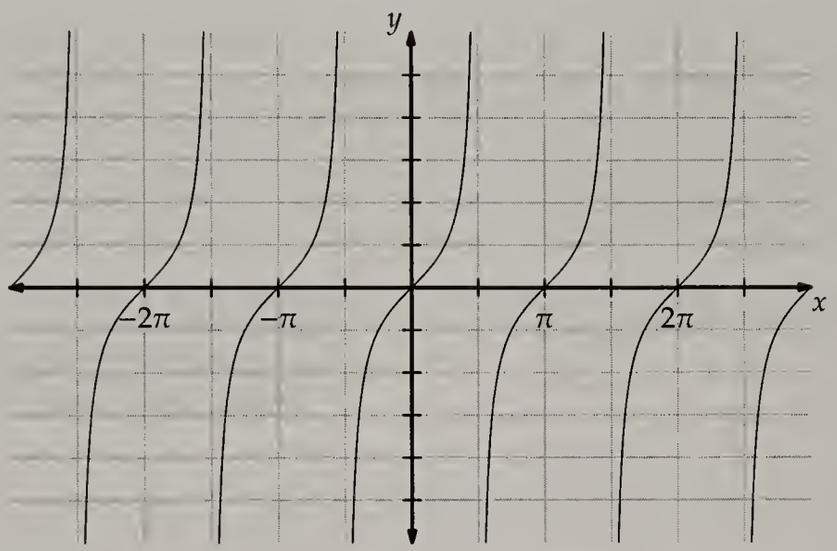


Figure 2.28: Graph of $y = \tan x$

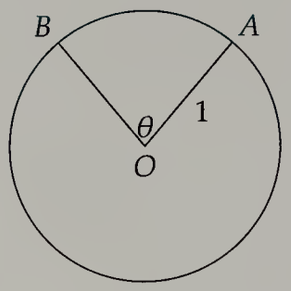
Three other trigonometric functions are **secant**, **cosecant**, and **cotangent**, which are denoted as sec, csc, and cot, respectively. These are defined in terms of cosine, sine, and tangent as follows:

$$\sec\theta = \frac{1}{\cos\theta},$$
$$\csc\theta = \frac{1}{\sin\theta},$$
$$\cot\theta = \frac{1}{\tan\theta} = \frac{\cos\theta}{\sin\theta}.$$

A trigonometric identity is an equation that is true for all angles for which the equation is defined. In this chapter, we proved the following identities:

$$\sin^2\theta + \cos^2\theta = 1,$$
$$\cos(-\theta) = \cos\theta,$$
$$\sin(-\theta) = -\sin\theta.$$

We also showed that $-1 \le \sin\theta \le 1$ and $-1 \le \cos\theta \le 1$ for all θ.



Degrees are not the only way we can measure angles. We can also use **radians**. To measure an angle in radians, we draw a circle with radius 1 centered at the vertex of the angle. The measure of the angle in radians equals the length of the arc cut off by the angle. So, in the diagram at right the measure of θ in radians equals the length of arc $\overset{\frown}{AB}$.

We say that $f(x)$ **varies sinusoidally** with a variable x, and that f is **sinusoidal**, if we can express f as $f(x) = a\sin(bx + c) + d$ for some constants a, b, c, and d.

A function f is called **periodic** if there is some nonzero constant t such that $f(x) = f(x+t)$ for all values of x such that $f(x)$ and $f(x + t)$ are defined. The smallest possible positive value of t is called the **period** of f (if such a smallest value exists). The period of $\sin ax$ and of $\cos ax$ is $\frac{2\pi}{|a|}$, while the period of $\tan ax$ is $\frac{\pi}{|a|}$.

The **amplitude** of a trig function is half the difference between the maximum and the minimum value of the function. The **phase shift** of a function of the form $f(x) = a\sin(bx+c)$, where a, b, and c are constants, is the smallest horizontal distance by which the graph of $g(x) = a\sin bx$ must be shifted to produce the graph of f. Typically, a positive phase shift indicates a rightward shift and a negative phase shift indicates a leftward shift.

By restricting the domains of the trig functions, we can define inverses of them. Below is a table of the inverse trig functions, their domains, and their ranges.

Trig function	Inverse Trig Function	Domain of Inverse	Range of Inverse
$\cos x$	$\arccos x$	$[-1, 1]$	$[0, \pi]$
$\sin x$	$\arcsin x$	$[-1, 1]$	$\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$
$\tan x$	$\arctan x$	$\mathbb{R}$	$\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$
$\cot x$	$\operatorname{arccot} x$	$\mathbb{R}$	$(0, \pi)$
$\sec x$	$\operatorname{arcsec} x$	$(-\infty, -1] \cup [1, +\infty)$	$\left[0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right]$
$\csc x$	$\operatorname{arccsc} x$	$(-\infty, -1] \cup [1, +\infty)$	$\left[-\frac{\pi}{2}, 0\right) \cup \left(0, \frac{\pi}{2}\right]$

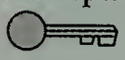
Things To Watch Out For!

WARNING!!

- In Section 1.3, we noted that $f^2(x)$ usually means $f(f(x))$. Trigonometric functions are a common exception to this. When we write $\sin^2 x$, we mean $(\sin x)^2$, not $\sin(\sin x)$.
- It's very easy to confuse the definitions of $\sec x$ and $\csc x$. Don't be afraid to look them up when you encounter secant and cosecant.
- If an angle measure does not include the degree symbol or any indication of units, then the measure is assumed to be in *radians*, not degrees. So, if you are asked to find $\sin 30$, the answer is not 1/2, since $\sin 30$ refers to the sine of an angle of 30 radians, not 30 degrees.
- When using a calculator or computer, make sure you know whether your device or program is using degrees or radians. If it assumes that the angle is in radians, and you enter "tan(30)" to find $\tan 30°$, the answer will not be what you intend. One quick way to test it is to use it to compute $\sin 30°$. If it returns something like -0.98803 instead of 0.5, then it is expecting the angles in radians, not degrees.

Problem Solving Strategies

Concepts:



- Graphs can help provide intuition for problems involving trig functions.
- Some problems involving inverse trig functions can be most easily solved by rewriting the problem in terms of the corresponding trig functions.
- When stuck on a trig problem involving tangent, cotangent, secant, and or cosecant, try rewriting the problem in terms of sine and cosine.
- Solving a simpler version of a complicated problem can help solve the original complex problem.
- In equations involving trig functions, try to isolate trig functions by manipulating the equation so that you have a trig function equal to a constant.
- Trig equations with a single type of trig function are often easier to deal with than equations with two different types of trig functions.

REVIEW PROBLEMS

2.39 Evaluate each of the following: (As an extra challenge, try finding them in your head, without writing anything down.)

(a) $\cos 150°$ (c) $\csc 90°$ (e) $\cos 315°$

(b) $\tan 30°$ (d) $\tan 270°$ (f) $\sin 225°$

2.40 Evaluate each of the following: (As an extra challenge, try finding them in your head, without writing anything down.)

(a) $\sin(-120°)$ (c) $\tan 1320°$

(b) $\cos(-540°)$ (d) $\sec 660°$

2.41 For how many values of θ such that $0 \leq \theta \leq 7\pi$ is $\cos\theta = -0.1$?

2.42 What are the domain and the range of the function $f(t) = 2\tan(3t - \pi)$?

2.43 Convert each of the following to radians:

(a) $144°$ (b) $336°$

2.44 Convert each of the following measures in radians to degrees:

(a) $\frac{5\pi}{2}$ (b) $\frac{11\pi}{3}$

2.45 Evaluate each of the following: (As an extra challenge, try finding these in your head, without writing anything down.)

(a) $\cos\frac{3\pi}{4}$ (c) $\cot\frac{5\pi}{4}$

(b) $\sin(-\pi)$ (d) $\sec\frac{7\pi}{6}$

2.46 Which of the following are positive: $\cos 1.3\pi$, $\sin(-6)$, $\tan \frac{7\pi}{5}$, $\sec 6$?

2.47 Determine the sign of $\sin 3 + \sin 4$ without using a calculator.

2.48 Is it possible for $\sin(\sin x)$ to equal 1?

2.49 Jenny is writing a computer program that contains the following line of code:

```
x = t*sin(y).
```

This line of code sets x equal to the product of t and the sine of y. Unfortunately, y is in degrees, and the "sin" function in her program expects it to be in radians. How should she change this line of code to obtain the desired result?

2.50

(a) What are the asymptotes of the graph of $y = \cot x$?

(b) What are the asymptotes of the graph of $y = \cot 3x$?

2.51 Michelle graphs a function of the form $f(x) = a \sin bx$, where a and b are nonzero constants, and then graphs $g(x) = a \cos bx$. She uses a blue pen for one graph and a red pen for the other. She gives you a 1-second glance at her two graphs—how can you quickly tell which color she used for the graph of f?

2.52 Find the period, amplitude, and phase shift of $g(x) = -2 \sin\left(\frac{x}{4} - \frac{\pi}{3}\right)$.

2.53 What is the period of $g(x) = -2 \tan(3x - \pi)$?

2.54

(a) Graph $y = \sin\left(\frac{\pi}{2} + x\right)$

(b) Graph $y = \sin\left(\frac{\pi}{2} - x\right)$.

(c) What relationship between $\sin\left(\frac{\pi}{2} - x\right)$ and $\cos x$ is suggested by the graph in part (b)?

2.55 Graph $y = 2 \sin\left(\frac{2x}{3} - \pi\right)$. (Be sure to check that you have the correct graph by choosing values of x, finding the corresponding values of y, and confirming that your graph passes through each ordered pair (x, y) you produce.)

2.56 Leah graphs the function $h(x) = a \cos bx$ for some values of a and b. The resulting graph is shown below. Find all possible values of a and b.

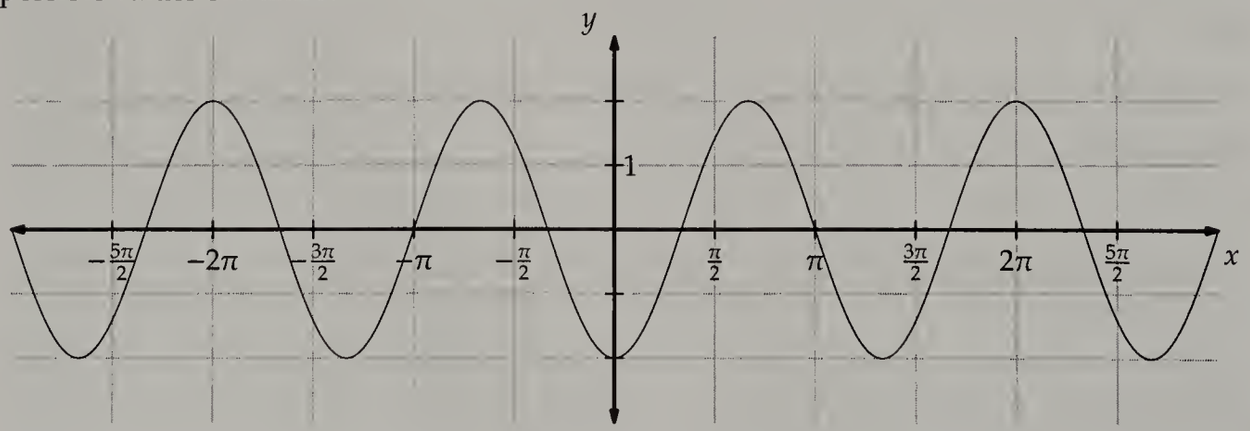


Figure 2.29: Graph for Problem 2.56

2.57 Evaluate each of the following:

(a) $\arccos(-0.5)$ (c) $\tan^{-1} 0$

(b) $\arcsin(-1)$ (d) $\arccos\left(\sqrt{3}/2\right)$

2.58 Evaluate $\sin(\arctan 2)$ without using a calculator.

2.59 Find all values of k such that $3\cot(4k - \pi) = \sqrt{3}$.

2.60 Graph $y = \arctan x$. What are the asymptotes of the graph?

2.61 The range of $\operatorname{arcsec} x$ is $\left[0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right]$. Why must we leave $\frac{\pi}{2}$ out of the range?

2.62 If $8\tan\theta = 3\cos\theta$, where $0 < \theta < \pi$, then determine the value of $\sin\theta$.

Challenge Problems

2.63

(a) How is the graph of $y = \sin\left(x - \frac{\pi}{2}\right)$ related to the graph of $y = \cos x$?

(b) What relationship between the functions $f(x) = \cos x$ and $g(x) = \sin\left(x - \frac{\pi}{2}\right)$ does your graph in part (a) suggest?

2.64 Steph graphs the function $h(x) = \sin(px + q)$ for some values of p and q. The resulting graph is shown below.

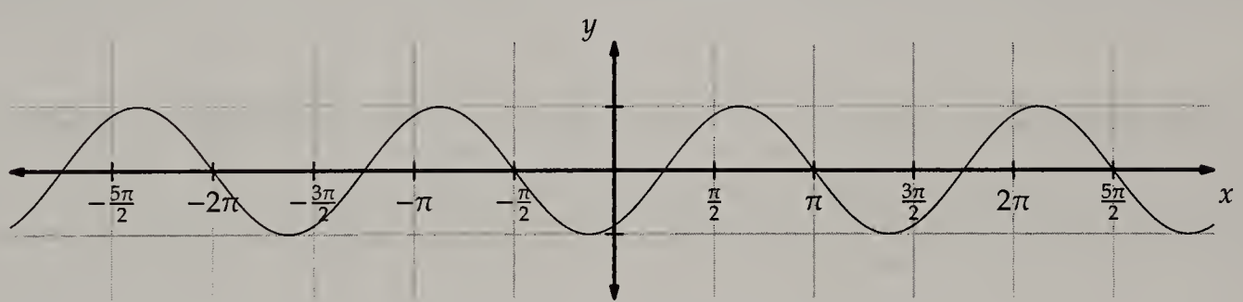


Figure 2.30: Graph for Problem 2.64

Find all possible values of p and q. **Hints:** 91

2.65 What is the period of $f(x) = (\sin x)(\cos x)$? **Hints:** 131, 239

2.66 How many values of x satisfy $\sin x = \frac{x^2}{625\pi^2}$? **Hints:** 237

2.67 Find $\arccos\left(\sin\frac{6\pi}{7}\right)$. **Hints:** 31, 187

2.68 For what values of θ is $\sin 3\theta < 0.5$? **Hints:** 158

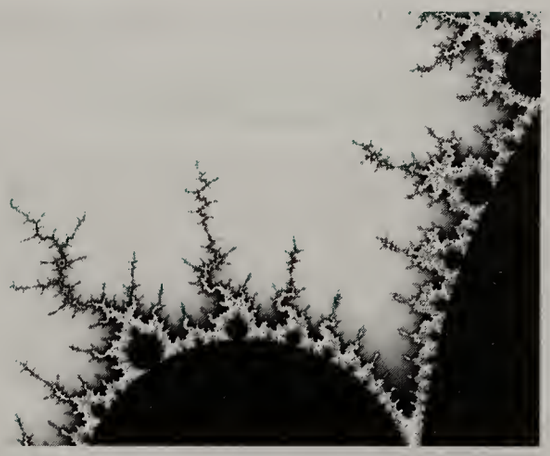
2.69 Let $g(x) = \sin ax + \cos bx$.

(a) If a and b are integers, must $g(x)$ be periodic?

(b) If a and b are rational numbers, must $g(x)$ be periodic?

2.70★ While finding the sine of a certain angle, an absent-minded professor failed to notice that his calculator was not in the correct angular mode. He was lucky to get the right answer. What are the two smallest positive values of x such that the sine of x degrees equals the sine of x radians? *(Source: AIME)* **Hints:** 83, 204, 55

2.71★ The price of Up-N-Down Stock varies sinusoidally with time. At the start of this year, the price of the stock is \$110 per share. Over the next 3 days, the price climbs to a peak of \$120 per share. If the lowest price the stock reaches is \$80 per share, how much money can you make this year (in 365 days starting from the start of this year) if you start with \$1000 and time all of your trades perfectly? (Assume that you can buy fractions of a share of stock, and that you can buy stock as many times as you want, but can never sell stock you haven't already bought first.) **Hints:** 193



Let it be an alliance of two large, formidable natures, mutually beheld, mutually feared before yet they recognize the deep identity which, beneath these disparities, unites them. – Ralph Waldo Emerson

CHAPTER 3

Trigonometric Identities

A **trigonometric identity** is an equation involving trig functions that is true for all values of the variables for which the equation is defined. Back on page 38 we proved one of the most important trig identities,

$$\sin^2\theta + \cos^2\theta = 1.$$

Throughout Chapter 2 we hinted at several other identities, and on page 50, we used the unit circle to prove $\sin(-\theta) = -\sin\theta$ and $\cos(-\theta) = \cos\theta$ for all θ. In this chapter, we derive and explore many more trig identities, and learn how to apply them to challenging problems. Don't try to memorize each new identity you encounter in this chapter. Instead, try to understand why the identities are true. Many of them will become automatic if you understand their derivations. Moreover, in Chapter 7, we'll see how complex numbers can help you remember many of these identities.

In this chapter, you will be asked to compute many trigonometric expressions. Do not use calculators or computers for these problems. The point of these problems is to learn how to apply identities to simplify expressions.

Important:

Don't bother memorizing all the identities in this chapter. We'll let you know which ones are the most important, and you should make sure you understand those well enough that you have them memorized. The main purpose of this chapter is to learn how deeply related the trigonometric functions are, and how to use those relationships to simplify expressions and solve problems. You can usually look up the more obscure identities whenever you need them. Focus instead on how to apply them.

In that vein, you should know the identities $\sin^2\theta + \cos^2\theta = 1$, $\cos(-\theta) = \cos\theta$, and $\sin(-\theta) = -\sin\theta$ well enough that you have them memorized. You can probably most easily internalize them by visualizing them on the unit circle.

3.1 Introduction to Trig Identities

Problems

Problem 3.1: In this problem, we find all x with $0 \le x < 2\pi$ such that $\tan x + \sec x = \sqrt{3}$.

(a) Write the equation in terms of $\sin x$ and $\cos x$.

(b) Solve the equation for $\sin x$ in terms of $\cos x$.

(c) Use $\sin^2 x + \cos^2 x = 1$ to find $\cos x$.

(d) Find all x with $0 \le x < 2\pi$ that satisfy the original equation. (Don't forget to check your answers.)

Problem 3.2: Prove the identity $\tan^2 x + 1 = \sec^2 x$.

Problem 3.3: Show that $\dfrac{(\csc x)(\sec x)}{1 + \tan^2 x} = \cot x$.

Problem 3.4: Use the unit circle to express each of the following in terms of $\sin\theta$, $\cos\theta$, or $\tan\theta$:

(a) $\sin(90° - \theta)$ (b) $\cos(90° - \theta)$ (c) $\tan(90° - \theta)$

Problem 3.5: Express each of the following in terms of $\sin\theta$, $\cos\theta$, $\tan\theta$, or $\cot\theta$:

(a) $\cos\left(\frac{\pi}{2} + \theta\right)$ (c) $\tan(\pi + \theta)$

(b) $\sin(\pi - \theta)$ (d) $\cos\left(\theta - \frac{3\pi}{2}\right)$

In each part, use the unit circle and/or graphs of trigonometric functions to explain your results.

Problem 3.6: In this problem, we find the smallest positive angle θ such that $\sin 2\theta = \cos 3\theta$.

(a) Use an identity from this section to write $\cos 3\theta$ in terms of sine. (You do not have to write it in terms of $\sin\theta$; you should express $\cos 3\theta$ as the sine of some expression.)

(b) Find an angle that satisfies your equation from part (a).

(c) Explain why the equation does not hold for any smaller positive value of θ.

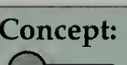
We'll start with an example that shows one reason why we care about trigonometric identities.

Problem 3.1: Find all x with $0 \le x < 2\pi$ such that $\tan x + \sec x = \sqrt{3}$.

Solution for Problem 3.1: We start by writing the equation in terms of sine and cosine.

Concept: A fruitful first step in many trig problems is rewriting the problem in terms of sine and cosine, because most people have more experience with sine and cosine than with the other trig functions.

Our equation now is $\dfrac{\sin x}{\cos x} + \dfrac{1}{\cos x} = \sqrt{3}$. Multiplying both sides by $\cos x$ gives $\sin x + 1 = \sqrt{3}\cos x$. Now what?



Concept: When solving an equation that contains trig functions, it's often best to write the equation in terms of a single trig function, so that you can then try to solve the equation for that function.

We write $\sin x + 1 = \sqrt{3}\cos x$ in terms of a single trig function using $\sin^2 x + \cos^2 x = 1$. Since $\cos^2 x = 1 - \sin^2 x$, we square both sides of the equation $\sin x + 1 = \sqrt{3}\cos x$, so that we can substitute $1 - \sin^2 x$ for the resulting $\cos^2 x$. Squaring the equation gives

$$(\sin x + 1)^2 = 3\cos^2 x.$$

Substituting $\cos^2 x = 1 - \sin^2 x$ and expanding both sides gives $\sin^2 x + 2\sin x + 1 = 3 - 3\sin^2 x$. Simplifying this equation gives $2\sin^2 x + \sin x - 1 = 0$. This equation is quadratic in $\sin x$. Factoring gives

$$(\sin x + 1)(2\sin x - 1) = 0.$$

From $\sin x + 1 = 0$, we have $\sin x = -1$, which gives us $x = \frac{3\pi}{2}$. (Remember, we are only asked for solutions such that $0 \le x < 2\pi$.) From $2\sin x - 1 = 0$, we have $\sin x = \frac{1}{2}$, which gives us the solutions $x = \frac{\pi}{6}$ and $x = \frac{5\pi}{6}$.

What's wrong with this finish:

Bogus Solution: Therefore, our solutions are $x = \frac{\pi}{6}$, $\frac{5\pi}{6}$, and $\frac{3\pi}{2}$.

If you try each of these "solutions," then you'll see a few problems. The value $x = \frac{\pi}{6}$ does indeed satisfy the original equation, but when $x = \frac{5\pi}{6}$, we have $\tan x = -\frac{\sqrt{3}}{3}$ and $\sec x = -\frac{2\sqrt{3}}{3}$, so $\tan x + \sec x = -\sqrt{3}$. Therefore, $x = \frac{5\pi}{6}$ is not a solution to the original equation. Moreover, when $x = \frac{3\pi}{2}$, both $\tan x$ and $\sec x$ are undefined. How did we get these false solutions?

WARNING!! Whenever you solve an equation by performing some possibly irreversible steps, you must check your solutions at the end, to make sure they satisfy the original equation. Two common irreversible steps are squaring an equation and multiplying both sides of an equation by an expression that might equal zero.

Here, we took both of the irreversible steps described in the warning; the first when we squared both sides and the second when we multiplied both sides by $\cos x$. The first of these ultimately brought us the solution $x = \frac{5\pi}{6}$ and the second brought us $x = \frac{3\pi}{2}$. Since these "solutions" do not satisfy the original equation, we call them **extraneous solutions**, and they are not solutions of the original equation. Therefore, the only solution to the original equation is $x = \frac{\pi}{6}$. □

In addition to helping us solve equations, the identity $\sin^2 x + \cos^2 x = 1$ can help us derive new identities.

Problem 3.2: Prove the identity $\tan^2 x + 1 = \sec^2 x$.

Solution for Problem 3.2: Writing the left side in terms of sine and cosine leads us quickly to the proof:

$$\tan^2 x + 1 = \frac{\sin^2 x}{\cos^2 x} + 1 = \frac{\sin^2 x + \cos^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = \sec^2 x.$$

□

As an Exercise, you'll also prove the related identity $\cot^2 x + 1 = \csc^2 x$.

Important:

$$\tan^2 x + 1 = \sec^2 x,$$
$$\cot^2 x + 1 = \csc^2 x.$$

Problem 3.3: Show that $\dfrac{(\csc x)(\sec x)}{1+\tan^2 x} = \cot x$.

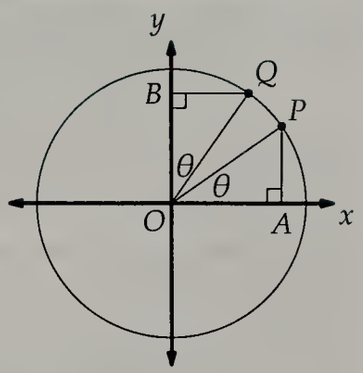
Solution for Problem 3.3: We could write everything in terms of sine and cosine, but we see that we can simplify the denominator of the left side immediately with $1 + \tan^2 x = \sec^2 x$. So, we do that first:

$$\frac{(\csc x)(\sec x)}{1+\tan^2 x} = \frac{(\csc x)(\sec x)}{\sec^2 x} = \frac{\csc x}{\sec x} = \frac{\frac{1}{\sin x}}{\frac{1}{\cos x}} = \frac{\cos x}{\sin x} = \cot x.$$

□

Problem 3.4: Use the unit circle to express each of $\sin(90° - \theta)$, $\cos(90° - \theta)$, and $\tan(90° - \theta)$ in terms of $\sin\theta$, $\cos\theta$, $\tan\theta$, or $\cot\theta$.

Solution for Problem 3.4: In the diagram, point P is the terminal point of the angle θ. To find the terminal point of $90° - \theta$, we start at $(0, 1)$, which is the terminal point of $90°$, and go θ *clockwise*, since θ is subtracted from $90°$, not added. If θ is acute, then drawing altitudes from P to the x-axis and Q to the y-axis produces right triangles $\triangle QBO$ and $\triangle PAO$. Since $QO = PO$ and $\angle QOB = \angle POA$, we have $\triangle QBO \cong \triangle PAO$. Therefore, $PA = QB$ and $OA = OB$.



Since P is the terminal point of θ, we have $\cos\theta = OA$ and $\sin\theta = PA$. Since Q is the terminal point of $90° - \theta$, we have $\cos(90° - \theta) = QB$ and $\sin(90° - \theta) = OB$. Combining these with $PA = QB$ and $OA = OB$ from above gives us

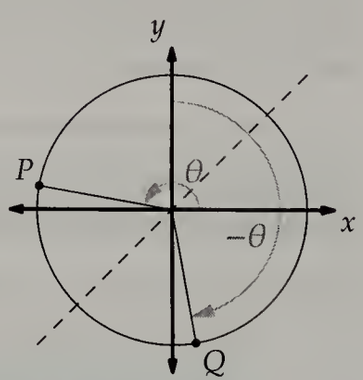
$$\cos(90° - \theta) = QB = PA = \sin\theta$$

and

$$\sin(90° - \theta) = OB = OA = \cos\theta.$$

We appear to have two new identities, but we haven't yet shown that these identities hold for all angles. We can extend our work above to other angles with some careful casework, or we can use the symmetry that our diagram suggests.

The actions of rotating the point $(1, 0)$ an angle θ counterclockwise and rotating $(0, 1)$ an angle θ clockwise are symmetric about the graph of $y = x$, which is dashed in the diagram at right. Therefore, the corresponding terminal points P and Q are symmetric about the line $y = x$. As we described on page 23, this means that we find the coordinates of Q by reversing the coordinates of P. Since P has coordinates $(\cos\theta, \sin\theta)$, point Q has coordinates $(\sin\theta, \cos\theta)$. But because Q is the terminal point of $90° - \theta$, we can also express Q as $(\cos(90° - \theta), \sin(90° - \theta))$. Since we can express Q both as $(\sin\theta, \cos\theta)$ and as $(\cos(90° - \theta), \sin(90° - \theta))$, we must have $\cos(90° - \theta) = \sin\theta$ and $\sin(90° - \theta) = \cos\theta$.



Important:

$$\sin(90° - \theta) = \cos\theta,$$
$$\cos(90° - \theta) = \sin\theta.$$

We can use these identities to find a similar one for tangent:

$$\tan(90° - \theta) = \frac{\sin(90° - \theta)}{\cos(90° - \theta)} = \frac{\cos\theta}{\sin\theta} = \cot\theta.$$

Important: $\tan(90° - \theta) = \cot\theta$

We could also have discovered the relationship between $\cos(90° - x)$ and $\sin x$ by examining the graphs of trigonometric functions. Converting to radians, we express $\cos(90° - x)$ as $\cos\left(\frac{\pi}{2} - x\right)$, and consider the graph of $y = \cos\left(\frac{\pi}{2} - x\right)$. To produce this graph, we first create the graph of $y = \cos\left(\frac{\pi}{2} + x\right)$ by shifting the graph of $y = \cos x$ to the left by $\frac{\pi}{2}$.

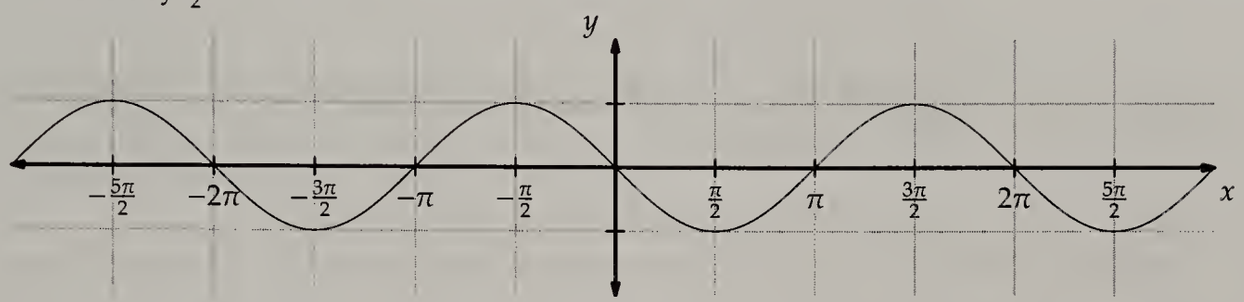


Figure 3.1: Graph of $y = \cos\left(\frac{\pi}{2} + x\right)$

We then produce the graph of $y = \cos\left(\frac{\pi}{2} - x\right)$ by reflecting the graph of $y = \cos\left(\frac{\pi}{2} + x\right)$ over the y-axis. (To see why, let $f(x) = \cos\left(\frac{\pi}{2} + x\right)$. Reflecting the graph of f over the y-axis gives us the graph of $y = f(-x) = \cos\left(\frac{\pi}{2} - x\right)$, as described on page 14.)

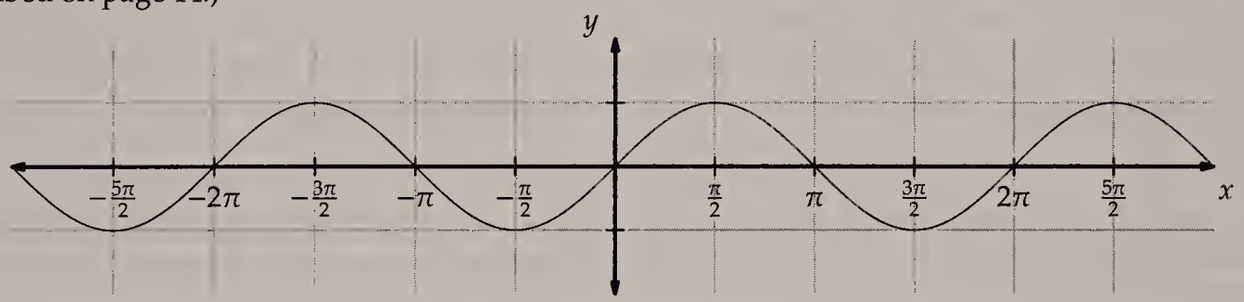


Figure 3.2: Graph of $y = \cos\left(\frac{\pi}{2} - x\right)$

The graph of $y = \cos\left(\frac{\pi}{2} - x\right)$ appears to be *exactly the same* as the graph of $y = \sin x$, which suggests that $\sin x = \cos\left(\frac{\pi}{2} - x\right)$ for all values of x. This graphical explanation is not a rigorous proof, but graphs can help us discover identities that we can rigorously prove by other means, such as our explanation with the unit circle. □

Problem 3.5: Express each of the following in terms of $\sin\theta$, $\cos\theta$, or $\tan\theta$:

(a) $\cos\left(\frac{\pi}{2} + \theta\right)$ (c) $\tan(\pi + \theta)$

(b) $\sin(\pi - \theta)$ (d) $\cos\left(\theta - \frac{3\pi}{2}\right)$

In each part, use the unit circle and/or graphs of trigonometric functions to explain your results.

Solution for Problem 3.5:

(a) What's wrong with this solution:

Bogus Solution: The graph of $y = \cos\left(\frac{\pi}{2} + x\right)$ is the graph of $y = \cos x$ shifted $\frac{\pi}{2}$ to the right. This graph is the same as the graph of $y = \sin x$, so we have $\cos\left(\frac{\pi}{2} + x\right) = \sin x$.

At first, we might question the claim that shifting the graph of $y = \cos x$ to the right by $\frac{\pi}{2}$ does indeed produce the graph of $y = \sin x$. So, let's try it:

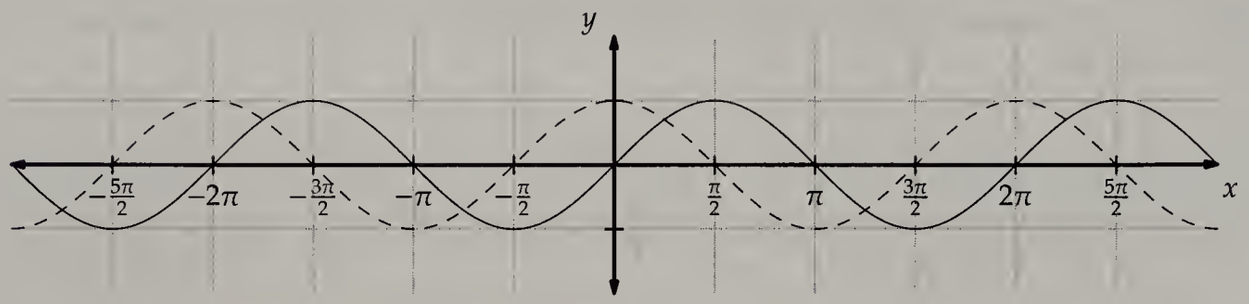


Figure 3.3: Graph of $y = \cos x$ (dashed) and $y = \cos x$ shifted $\frac{\pi}{2}$ to the right (solid)

Sure enough, the resulting solid graph above is the graph of $y = \sin x$. Maybe the Bogus Solution is correct.

Important: When you think you've found a new trigonometric identity, test it for several angles. Make sure you use angles from different quadrants in your testing—a proposed identity might work only in a few special cases.

We test whether $\cos\left(\frac{\pi}{2} + \theta\right)$ always equals $\sin\theta$ by evaluating both for various values of θ. Let's try $\theta = 0$. That gives us $\cos\left(\frac{\pi}{2} + \theta\right) = 0$ and $\sin\theta = 0$. OK, that works; let's try $\theta = \frac{\pi}{4}$. Then, we have $\cos\left(\frac{\pi}{2} + \theta\right) = \cos\frac{3\pi}{4} = -\frac{\sqrt{2}}{2}$ and $\sin\theta = \frac{\sqrt{2}}{2}$. Uh-oh. Those aren't equal; they're opposites. If we try a few more values of θ, it appears that $\cos\left(\frac{\pi}{2} + \theta\right)$ and $\sin\theta$ are always opposites.

Looking more closely at our Bogus Solution reveals the error. The graph of $y = \cos\left(\frac{\pi}{2} + x\right)$ is the result of shifting the graph of $y = \cos x$ to the *left* by $\frac{\pi}{2}$, not to the right. (See page 13 if you don't see why.) The correct graph is below:

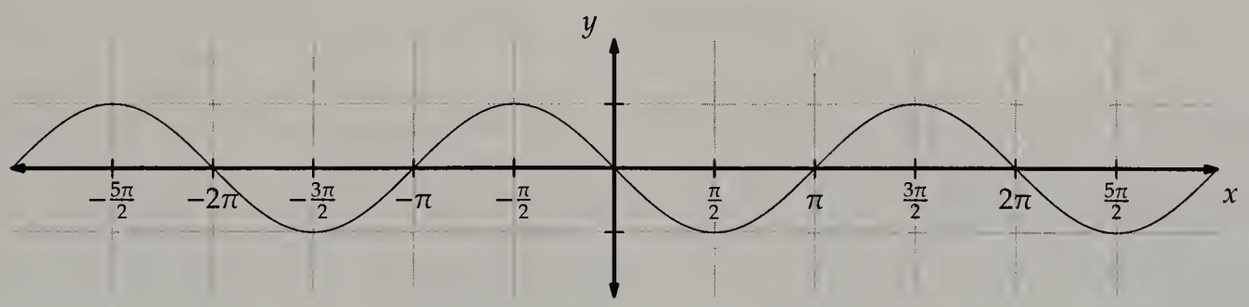
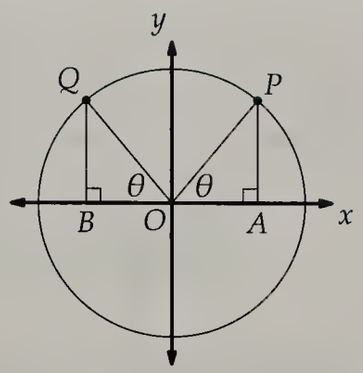


Figure 3.4: Graph of $y = \cos\left(\frac{\pi}{2} + x\right)$

This graph is the same as the graph of $y = \sin x$ reflected over the x-axis. Reflecting the graph of $y = \sin x$ over the x-axis gives us the graph of $y = -\sin x$ (see page 14). Since the graphs of $y = \cos\left(\frac{\pi}{2} + x\right)$ and $y = -\sin x$ appear to be the same, it seems like we must have $\cos\left(\frac{\pi}{2} + x\right) = -\sin x$ for all x.

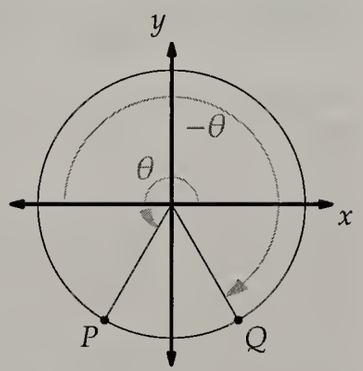
See if you can also use the unit circle to explain this identity.

(b) We're asked to find $\sin(\pi-\theta)$ in terms of trigonometric functions of θ. Suppose θ is an acute angle, and that P is its terminal point. Let Q be the terminal point of $\pi - \theta$, so that the segment joining Q with the origin forms an angle of θ with the negative x-axis, as shown at right. Drawing altitudes from P and Q to the x-axis forms congruent right triangles POA and QOB, from which we have $PA = QB$. Therefore, P and Q are the same distance above the x-axis, so they have the same y-coordinates, which means their corresponding angles have the same sine.



If θ is obtuse ($\frac{\pi}{2} < \theta < \pi$), then we have the essentially the same diagram, but with P to the left of the y-axis and Q to the right of the y-axis.

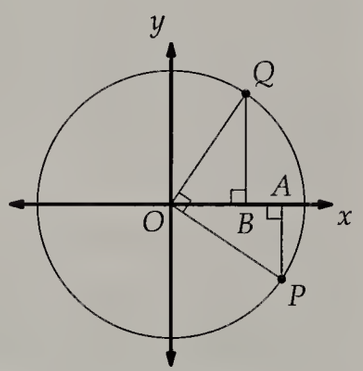
We can go through similar arguments for values of θ greater than π or less than 0. We could also note that rotating the point $(1,0)$ an angle of θ degrees counterclockwise and rotating the point $(-1,0)$ an angle of θ degrees clockwise are symmetric about the y-axis, as depicted at right.



If two points are symmetric about the y-axis, then the points have opposite x-coordinates and equal y-coordinates. Therefore, we have $\cos(\pi-\theta) = -\cos\theta$ and $\sin(\pi - \theta) = \sin\theta$.

(c) Back on page 50, we showed that the period of the tangent function is π. Therefore, we have $\tan(\pi + \theta) = \tan\theta$.

(e) We must find $\cos\left(\theta - \frac{3\pi}{2}\right)$ in terms of trigonometric functions of θ. At right, we show a case in which θ is a fourth quadrant angle (with terminal point P), so that $\theta - \frac{3\pi}{2}$ is a first quadrant angle (with terminal point Q). We draw altitudes $\overline{PA}$ and $\overline{QB}$ as shown. Since P and Q are the terminal points of angles that are $\frac{3\pi}{2}$ apart, we have $\angle POQ = \frac{\pi}{2}$, which means that $\angle QOB = \frac{\pi}{2} - \angle POA = \angle OPA$. This tells us that $\triangle QOB \cong \triangle OPA$. Therefore, we have $OB = PA$.



Since P and Q are the terminal points of θ and $\theta - \frac{3\pi}{2}$, respectively, we have $\sin\theta = -PA$ (P is below the x-axis, so $\sin\theta$ is negative) and $\cos\left(\theta - \frac{3\pi}{2}\right) = OB$. Combining these with $OB = PA$ gives us $\cos\left(\theta - \frac{3\pi}{2}\right) = -\sin\theta$.

We can also visualize this identity by graphing $y = \cos\left(x - \frac{3\pi}{2}\right)$, which we do by shifting the graph of $y = \cos x$ to the right by $\frac{3\pi}{2}$. The result is shown below:

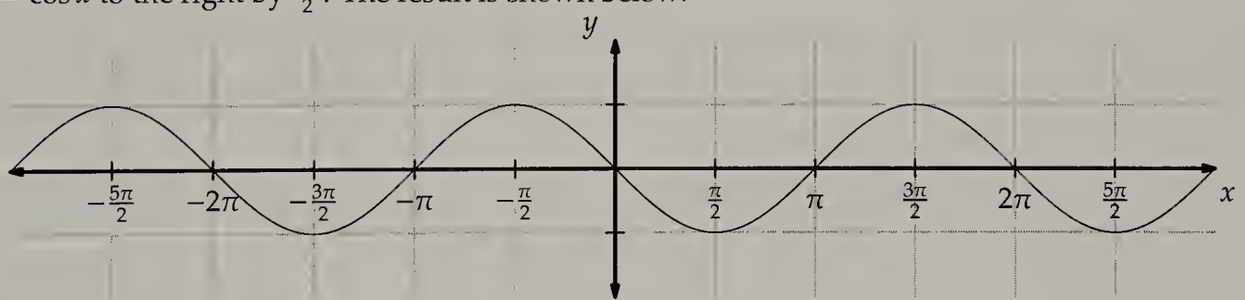


Figure 3.5: Graph of $y = \cos\left(x - \frac{3\pi}{2}\right)$

Just as in part (a), the resulting graph is the same as the graph of $y = \sin x$ reflected over the x-axis. We deduce that the graphs of $y = \cos\left(x - \frac{3\pi}{2}\right)$ and $y = -\sin x$ are the same, so $\cos\left(x - \frac{3\pi}{2}\right) = -\sin x$ for all x.

Testing with $x = \frac{7\pi}{4}$ gives us $\cos\left(x - \frac{3\pi}{2}\right) = \cos\frac{\pi}{4} = \frac{\sqrt{2}}{2}$ and $-\sin x = -\left(-\frac{\sqrt{2}}{2}\right) = \frac{\sqrt{2}}{2}$. Testing with $x = \frac{\pi}{6}$ gives us $\cos\left(x - \frac{3\pi}{2}\right) = \cos\left(-\frac{4\pi}{3}\right) = -\frac{1}{2}$ and $-\sin x = -\frac{1}{2}$. As expected, both our tests work. □

Don't bother memorizing all the possible identities we can form by shifting an angle that is some multiple of 90°. As we just saw, thinking about the unit circle or about graphs of trigonometric functions can help you rediscover these identities whenever you need them. Probably the most commonly used are the ones we showed in Problem 3.4 for functions of $90° - \theta$ and those for functions of $180° - \theta$:

Important:

$$\sin(180° - \theta) = \sin\theta,$$
$$\cos(180° - \theta) = -\cos\theta.$$

Problem 3.6: Find the smallest positive angle θ such that $\sin 2\theta = \cos 3\theta$.

Solution for Problem 3.6: We don't have good tools yet for expressing $\sin 2\theta$ and $\cos 3\theta$ in terms of $\sin\theta$ and $\cos\theta$. However, we can turn $\cos 3\theta$ into the sine of an expression by using the identity $\cos x = \sin(90° - x)$, which turns $\sin 2\theta = \cos 3\theta$ into

$$\sin 2\theta = \sin(90° - 3\theta).$$

Now, we at least see a way to find one solution; the sines of two angles are equal if the angles are equal. Solving $2\theta = 90° - 3\theta$ gives us $\theta = 18°$. But how do we know that this is the smallest possible positive solution?

Consider what happens to both sides of the equation if we have $0° < \theta < 18°$. For $\theta = 18°$, we found that we have $\sin 2\theta = \cos 3\theta$, or $\sin 36° = \cos 54°$. When we decrease θ from 18°, the value of $\sin 2\theta$ decreases, because as 2θ decreases from 36° (but stays positive), the terminal point of 2θ gets closer to $(1, 0)$. That is, it gets closer to the x-axis, which means that $\sin 2\theta$ decreases as 2θ decreases from 36° to 0°. Meanwhile, as θ decreases from 18° to 0°, the value of $\cos 3\theta$ *increases*, since the terminal point of 3θ gets farther to the right of the y-axis as 3θ decreases from 54° to 0°. Putting these observations together with $\sin 36° = \cos 54°$, we conclude that if $0° < \theta < 18°$, then we have

$$\sin 2\theta < \sin 36° = \cos 54° < \cos 3\theta.$$

Therefore, we cannot have $\sin 2\theta = \cos 3\theta$ for any positive value of θ less than 18°. □

Exercises

3.1.1 Prove the identity $\cot^2\theta + 1 = \csc^2\theta$.

3.1.2 Prove that $\sin(360° - \theta) = -\sin\theta$ for all angles θ.

3.1.3 Use the identity you proved in Problem 3.2 to solve Problem 3.1 without first converting the equation to sines and cosines.

3.1.4 Find $\cos\left(\frac{3\pi}{2} + \theta\right)$ in terms of $\sin\theta$ or $\cos\theta$.

3.1.5 Find the value of the product $\tan\left(\frac{\pi}{12}\right)\tan\left(\frac{2\pi}{12}\right)\tan\left(\frac{3\pi}{12}\right)\tan\left(\frac{4\pi}{12}\right)\tan\left(\frac{5\pi}{12}\right)$.

3.1.6 Find the range of

$$f(A) = \frac{\sin A(3\cos^2 A + \cos^4 A + 3\sin^2 A + \sin^2 A\cos^2 A)}{\tan A(\sec A - \sin A\tan A)},$$

where A is not an integer multiple of $\pi/2$. *(Source: HMMT)*

3.2 Sums and Differences of Angles

In the previous section, we found identities to simplify expressions like $\sin\left(\frac{\pi}{2} + \theta\right)$ and $\cos(\theta - 180°)$. In each of the expressions we simplified, one of the angles was a multiple of 90°. In this section, we develop identities for trigonometric functions of the sum or difference of any two angles.

We start by proving that

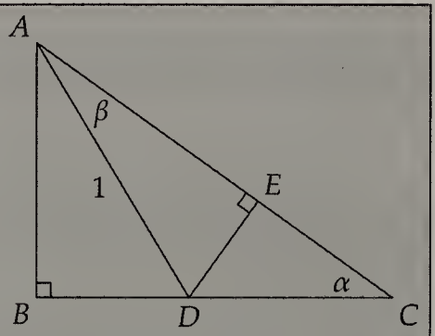
$$\sin(\alpha + \beta) = \sin\alpha\cos\beta + \sin\beta\cos\alpha$$

for positive angles α and β such that $\alpha + \beta < 90°$. We can extend this proof to all angles with a lot of careful casework; we won't do so here. (You'll have a chance to cover some of these cases as an Exercise.) You can use this identity for all problems after Problem 3.7.

Problems

Problem 3.7: In this problem, we will use the diagram at right to prove that if α and β are positive angles such that $\alpha + \beta < 90°$, then

$$\sin(\alpha + \beta) = \sin\alpha\cos\beta + \sin\beta\cos\alpha$$



In the diagram, we have $AD = 1$, $\angle BCA = \alpha$, and $\angle DAE = \beta$.

(a) Explain why $AE = \cos\beta$ and $DE = \sin\beta$.

(b) Find EC in terms of sine and cosine of α and β.

(c) Find $\angle ADB$ in terms of α and β.

(d) Show that $AB = \sin(\alpha + \beta)$ and $\sin\alpha = \frac{AB}{AC}$.

(e) Show that $\sin(\alpha + \beta) = \sin\alpha\cos\beta + \sin\beta\cos\alpha$.

Problem 3.8: Use the identity you found in Problem 3.7 together with identities from the previous section to find $\sin(\alpha - \beta)$ in terms of $\sin\alpha$, $\cos\alpha$, $\sin\beta$, and $\cos\beta$. Test your identity with $\alpha = 120°$ and $\beta = 60°$.

Problem 3.9: In Problem 3.4, we found that $\sin(90° - \alpha) = \cos\alpha$.

(a) Use $\sin(90° - \alpha) = \cos\alpha$ together with the identity for $\sin(\alpha + \beta)$ to express $\cos(\alpha - \beta)$ in terms of $\sin\alpha$, $\cos\alpha$, $\sin\beta$ and $\cos\beta$.

(b) Test your identity in the previous part by letting $\alpha = 120°$ and $\beta = 60°$. If the result is not a true statement, then try the first part again.

(c) Find an identity for $\cos(\alpha + \beta)$.

Problem 3.10: Suppose that $\tan\alpha$, $\tan\beta$, and $\tan(\alpha+\beta)$ are all defined. Find $\tan(\alpha+\beta)$ in terms of $\tan\alpha$ and $\tan\beta$.

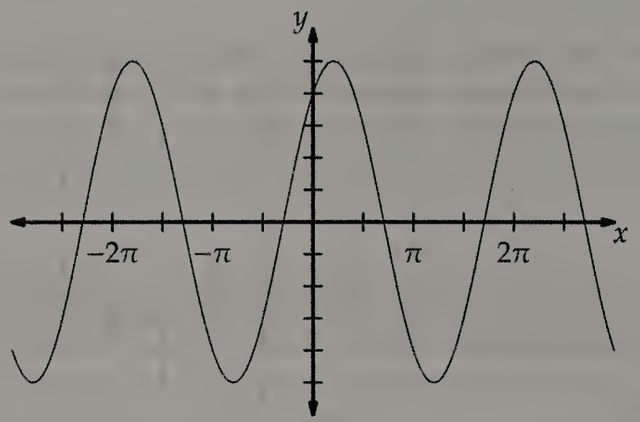
Problem 3.11:

(a) Find $\cos 15°$.

(b) Find $\sin\frac{13\pi}{12}$.

Problem 3.12: Find the value of $\cos\left(\arcsin\frac{4}{5} + \arctan\frac{5}{12}\right)$ without using a calculator.

Problem 3.13: Below is the graph of $y = 3\sin x + 4\cos x$:



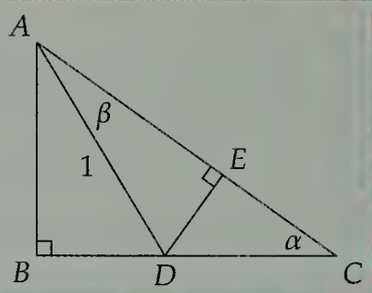
It looks like the graph of a function of the form $a\sin(bx + c)$! Why?

Problem 3.14: Compute x if $\arctan x + \arctan 1 = 2(\arctan x - \arctan \frac{1}{3})$. *(Source: ARML)*

Problem 3.7: Use the diagram at right to prove that if α and β are positive angles such that $\alpha + \beta < 90°$, then

$$\sin(\alpha + \beta) = \sin\alpha\cos\beta + \sin\beta\cos\alpha$$

In the diagram, we have $AD = 1$, $\angle BCA = \alpha$, and $\angle DAE = \beta$.



Solution for Problem 3.7: We start by finding the lengths of various segments in the diagram in terms of sine and cosine of α and β. Applying the right triangle relationships we learned on page 35 to $\triangle ADE$, we have $\cos\beta = \frac{AE}{AD}$ and $\sin\beta = \frac{DE}{AD}$. We therefore have $AE = \cos\beta$ and $DE = \sin\beta$.

Turning to $\triangle DEC$, we have $\tan\alpha = \frac{DE}{EC}$, so

$$EC = \frac{DE}{\tan\alpha} = \frac{\sin\beta\cos\alpha}{\sin\alpha}.$$

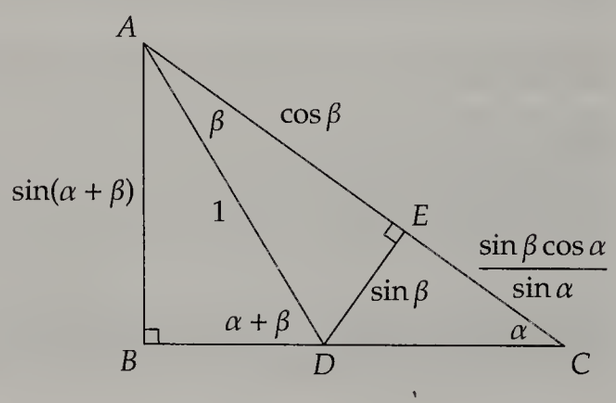
Next, we consider right triangle ADB. To find either leg of this triangle in terms of trig functions of α and β, we'll first have to find the acute angles of $\triangle ADB$ in terms of α and β. From the angles in $\triangle ADC$, we have $\angle ADC = 180° - \angle DAC - \angle DCA = 180° - \alpha - \beta$. Therefore, we have

$$\angle ADB = 180° - \angle ADC = 180° - (180° - \alpha - \beta) = \alpha + \beta.$$

We can now use $\triangle ABD$ to find an expression for $\sin(\alpha + \beta)$:

$$\sin(\alpha + \beta) = \sin\angle ADB = \frac{AB}{AD}.$$

Since $AD = 1$, we have $AB = \sin(\alpha + \beta)$, as shown in the diagram.



We're almost there! We still have one more right triangle to investigate: $\triangle ABC$. This triangle gives us

$$\sin\alpha = \sin\angle ACB = \frac{AB}{AC} = \frac{AB}{AE+EC} = \frac{\sin(\alpha+\beta)}{\cos\beta + \frac{\sin\beta\cos\alpha}{\sin\alpha}},$$

so we have the equation

$$\frac{\sin(\alpha+\beta)}{\cos\beta + \frac{\sin\beta\cos\alpha}{\sin\alpha}} = \sin\alpha.$$

Multiplying both sides by the denominator of the left side gives:

$$\begin{aligned}\sin(\alpha+\beta) &= \sin\alpha\left(\cos\beta + \frac{\sin\beta\cos\alpha}{\sin\alpha}\right)\\ &= \sin\alpha\cos\beta + \sin\beta\cos\alpha,\end{aligned}$$

as desired. □

With a great deal of careful casework, we can prove that this identity holds for any angles α and β. You'll see some of these cases as an Exercise. We can now use this identity to prove other identities for trigonometric functions of sums or differences of angles.

Concept: Once we find one trigonometric identity, we can often use it to quickly prove other related identities.

Problem 3.8: Use the identity you found in Problem 3.7 together with identities from the previous section to find $\sin(\alpha-\beta)$ in terms of $\sin\alpha$, $\cos\alpha$, $\sin\beta$, and $\cos\beta$. Test your identity with $\alpha = 120°$ and $\beta = 60°$.

Solution for Problem 3.8: We write $\sin(\alpha-\beta)$ as $\sin(\alpha+(-\beta))$ and use the identity for the sine of a sum of angles:

$$\sin(\alpha-\beta) = \sin(\alpha+(-\beta)) = \sin\alpha\cos(-\beta) + \sin(-\beta)\cos\alpha.$$

Now, we apply $\sin(-\beta) = -\sin\beta$ and $\cos(-\beta) = \cos\beta$ to the right side to find:

$$\sin(\alpha-\beta) = \sin\alpha\cos\beta - \sin\beta\cos\alpha.$$

If we let $\alpha = 120°$ and $\beta = 60°$, we have $\sin(\alpha-\beta) = \sin 60° = \frac{\sqrt{3}}{2}$ and

$$\sin\alpha\cos\beta - \sin\beta\cos\alpha = \sin 120°\cos 60° - \sin 60°\cos 120° = \left(\frac{\sqrt{3}}{2}\right)\left(\frac{1}{2}\right) - \left(\frac{\sqrt{3}}{2}\right)\left(-\frac{1}{2}\right) = \frac{\sqrt{3}}{2}.$$

□

Now, we turn to cosine.

Problem 3.9: Find identities for $\cos(\alpha-\beta)$ and $\cos(\alpha+\beta)$.

Solution for Problem 3.9: We already have angle sum and difference identities for sine, so we look for ways to relate sine and cosine. Two identities come to mind, $\sin^2\theta + \cos^2\theta = 1$ and $\sin(90°-\theta) = \cos\theta$. The first doesn't look immediately helpful, because squaring the identities we already have for $\sin(\alpha+\beta)$ and $\sin(\alpha-\beta)$ looks scary. So, we focus instead on $\sin(90°-\theta) = \cos\theta$, which gives us

$$\cos(\alpha-\beta) = \sin(90°-(\alpha-\beta)) = \sin(90°-\alpha+\beta).$$

Now we have the sine of a combination of three angles. By grouping the first two, we can write it as the sine of a sum of two angles, which we know how to handle:

$$\cos(\alpha - \beta) = \sin(90^\circ - \alpha + \beta) = \sin((90^\circ - \alpha) + \beta).$$

We can now apply the identity

$$\sin(x + y) = \sin x \cos y + \sin y \cos x.$$

By letting $x = 90^\circ - \alpha$ and $y = \beta$, we have

$$\cos(\alpha - \beta) = \sin((90^\circ - \alpha) + \beta) = \sin(90^\circ - \alpha)\cos\beta + \sin\beta\cos(90^\circ - \alpha).$$

Finally, since $\sin(90^\circ - \alpha) = \cos\alpha$ and $\cos(90^\circ - \alpha) = \sin\alpha$, we have

$$\cos(\alpha - \beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta.$$

We can either do essentially the same thing for $\cos(\alpha + \beta)$, or we can write $\cos(\alpha + \beta)$ as $\cos(\alpha - (-\beta))$ and use the identity we just proved:

$$\begin{aligned}\cos(\alpha + \beta) &= \cos(\alpha - (-\beta)) \\ &= \cos\alpha\cos(-\beta) + \sin\alpha\sin(-\beta) \\ &= \cos\alpha\cos\beta - \sin\alpha\sin\beta,\end{aligned}$$

where we used $\cos(-\beta) = \cos\beta$ and $\sin(-\beta) = -\sin\beta$ in the last step. □

Finally, we tackle tangent.

Problem 3.10: Suppose that $\tan\alpha$, $\tan\beta$, and $\tan(\alpha + \beta)$ are all defined. Find $\tan(\alpha + \beta)$ in terms of $\tan\alpha$ and $\tan\beta$.

Solution for Problem 3.10: We already know how to handle sine and cosine, so we write tangent in terms of sine and cosine, and use the identities we've already proved:

$$\tan(\alpha + \beta) = \frac{\sin(\alpha + \beta)}{\cos(\alpha + \beta)} = \frac{\sin\alpha\cos\beta + \sin\beta\cos\alpha}{\cos\alpha\cos\beta - \sin\alpha\sin\beta}.$$

Now what? We want an identity in terms of $\tan\alpha$ and $\tan\beta$, but we just have sines and cosines.

There are a few ways we can reach our desired identity. We could divide the numerator and denominator of the our huge fraction above by $\cos\alpha\cos\beta$ (make sure you see why this works), or we can note that since $\tan\alpha = \frac{\sin\alpha}{\cos\alpha}$, we have $\sin\alpha = \tan\alpha\cos\alpha$. Similarly, we have $\sin\beta = \tan\beta\cos\beta$, and we can write our identity as:

$$\begin{aligned}\tan(\alpha + \beta) &= \frac{\sin\alpha\cos\beta + \sin\beta\cos\alpha}{\cos\alpha\cos\beta - \sin\alpha\sin\beta} \\ &= \frac{\tan\alpha\cos\alpha\cos\beta + \tan\beta\cos\beta\cos\alpha}{\cos\alpha\cos\beta - \tan\alpha\cos\alpha\tan\beta\cos\beta} \\ &= \frac{(\cos\alpha\cos\beta)(\tan\alpha + \tan\beta)}{(\cos\alpha\cos\beta)(1 - \tan\alpha\tan\beta)}.\end{aligned}$$

Because $\tan\alpha$ and $\tan\beta$ are both defined, we know that $\cos\alpha\cos\beta \neq 0$. Therefore, the $\cos\alpha\cos\beta$ terms in the numerator and denominator cancel and leave

$$\tan(\alpha + \beta) = \frac{\tan\alpha + \tan\beta}{1 - \tan\alpha\tan\beta}.$$

□

As an Exercise, you'll prove the corresponding identity for $\tan(\alpha - \beta)$.

WARNING!! The identity

$$\tan(\alpha+\beta)=\frac{\tan\alpha+\tan\beta}{1-\tan\alpha\tan\beta}$$

only holds when $1-\tan\alpha\tan\beta\neq 0$ and when both $\tan\alpha$ and $\tan\beta$ are defined (that is, when $\cos\alpha\neq 0$ and $\cos\beta\neq 0$). If $\tan\alpha\tan\beta=1$, then $\tan(\alpha+\beta)$ is undefined.

Let's take a closer look at why $\tan(\alpha+\beta)$ is undefined when $\tan\alpha\tan\beta=1$. From $\tan\alpha\tan\beta=1$, we have $\sin\alpha\sin\beta=\cos\alpha\cos\beta$, so $\cos\alpha\cos\beta-\sin\alpha\sin\beta=0$. The left side equals $\cos(\alpha+\beta)$, which means $\cos(\alpha+\beta)=0$, so $\tan(\alpha+\beta)$ is undefined.

We now have identities for the sine, cosine, and tangent of a sum or difference of two angles:

Important: **Angle sum and difference identities:**

$$\begin{aligned}
\sin(\alpha+\beta)&=\sin\alpha\cos\beta+\sin\beta\cos\alpha,\\
\sin(\alpha-\beta)&=\sin\alpha\cos\beta-\sin\beta\cos\alpha,\\
\cos(\alpha+\beta)&=\cos\alpha\cos\beta-\sin\alpha\sin\beta,\\
\cos(\alpha-\beta)&=\cos\alpha\cos\beta+\sin\alpha\sin\beta,\\
\tan(\alpha+\beta)&=\frac{\tan\alpha+\tan\beta}{1-\tan\alpha\tan\beta},\\
\tan(\alpha-\beta)&=\frac{\tan\alpha-\tan\beta}{1+\tan\alpha\tan\beta}.
\end{aligned}$$

These identities are sometimes expressed in more concise form using $\pm$ and $\mp$:

$$\begin{aligned}
\sin(\alpha\pm\beta)&=\sin\alpha\cos\beta\pm\sin\beta\cos\alpha,\\
\cos(\alpha\pm\beta)&=\cos\alpha\cos\beta\mp\sin\alpha\sin\beta,\\
\tan(\alpha\pm\beta)&=\frac{\tan\alpha\pm\tan\beta}{1\mp\tan\alpha\tan\beta}.
\end{aligned}$$

The $\pm$ symbols on the right sides mean "match the sign of the $\pm$ on the left," while the $\mp$ symbols on the right sides mean "take the opposite of the sign of the $\pm$ on the left."

Don't bother memorizing all of these now. As you've just seen, if you only have $\sin(\alpha+\beta)$ memorized, you can derive all the rest pretty quickly. Moreover, in Chapter 7, we'll learn an even faster way to recall these identities.

Sidenote: If you aren't quite sure you have remembered an identity correctly, you can easily test your memory by choosing easy angles to substitute into your identity. For example, if you aren't sure whether $\sin(\alpha+\beta)$ equals $\sin\alpha\cos\beta+\sin\beta\cos\alpha$ or $\sin\alpha\cos\beta-\sin\beta\cos\alpha$, then letting $\alpha=0°$ and $\beta=90°$ in both will quickly rule out the latter.

Problem 3.11:

(a) Find $\cos 15°$.

(b) Find $\sin\frac{13\pi}{12}$.

Solution for Problem 3.11:

(a) If we start in the usual way, by building a right triangle on the unit circle, we end up with a right triangle with acute angles with measures 15° and 75° and hypotenuse length 1. Unfortunately, we don't know anything (yet!) about the sides of such a triangle. However, we might be able to use our new angle sum and difference identities to express 15° in terms of angles whose sine and cosine we know. We list a few of the angles we know how to handle: 0°, 30°, 45°, 60°, 90°. Aha! We have $45° - 30° = 15°$, so we can use our difference of angles formula:

$$\begin{aligned}\cos 15° &= \cos(45° - 30°) \\ &= \cos 45° \cos 30° + \sin 45° \sin 30° \\ &= \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{3}}{2} + \frac{\sqrt{2}}{2} \cdot \frac{1}{2} \\ &= \frac{\sqrt{6} + \sqrt{2}}{4}.\end{aligned}$$

(b) We know that adding fractions of the form $\frac{a}{3}$ and $\frac{b}{4}$ can give us a fraction with 12 as the denominator, and we know how to handle sine and cosine of angles like $\frac{\pi}{4}$ and $\frac{\pi}{3}$. So, again, we look over the angles we know how to handle, writing them with denominator 12, and try to find a pair we can combine to get $\frac{13\pi}{12}$: $\frac{3\pi}{12}$, $\frac{4\pi}{12}$, $\frac{6\pi}{12}$, $\frac{8\pi}{12}$, $\frac{9\pi}{12}$. Success! We have $\frac{4\pi}{12} + \frac{9\pi}{12} = \frac{13\pi}{12}$, so:

$$\begin{aligned}\sin \frac{13\pi}{12} &= \sin\left(\frac{9\pi}{12} + \frac{4\pi}{12}\right) = \sin\left(\frac{3\pi}{4} + \frac{\pi}{3}\right) \\ &= \sin \frac{3\pi}{4} \cos \frac{\pi}{3} + \sin \frac{\pi}{3} \cos \frac{3\pi}{4} \\ &= \frac{\sqrt{2}}{2} \cdot \frac{1}{2} + \frac{\sqrt{3}}{2} \cdot \left(-\frac{\sqrt{2}}{2}\right) \\ &= \frac{\sqrt{2} - \sqrt{6}}{4}.\end{aligned}$$

As a quick check, we can note that this answer is negative, which is what we should expect, since the sine of a third quadrant angle must be negative.

□

Problem 3.12: Find the value of $\cos\left(\arcsin \frac{4}{5} + \arctan \frac{5}{12}\right)$ without using a calculator.

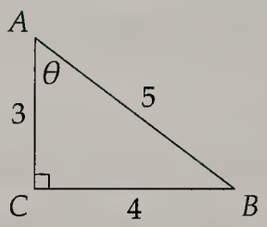
Solution for Problem 3.12: If we could calculate the angles that equal $\arcsin \frac{4}{5}$ and $\arctan \frac{5}{12}$, we could simply add these angles and take the cosine of the result. Unfortunately, we can't find these two angles without a calculator. However, we saw in Section 2.5 that we can calculate expressions like $\cos\left(\arcsin \frac{4}{5}\right)$, and we can use the angle sum formula to break $\cos\left(\arcsin \frac{4}{5} + \arctan \frac{5}{12}\right)$ into expressions like $\cos\left(\arcsin \frac{4}{5}\right)$:

$$\cos\left(\arcsin \frac{4}{5} + \arctan \frac{5}{12}\right) = \cos\left(\arcsin \frac{4}{5}\right) \cos\left(\arctan \frac{5}{12}\right) - \sin\left(\arcsin \frac{4}{5}\right) \sin\left(\arctan \frac{5}{12}\right).$$

One of these expressions is pretty easy: by the definition of arcsin, we have $\sin\left(\arcsin \frac{4}{5}\right) = \frac{4}{5}$. Let's tackle $\cos\left(\arcsin \frac{4}{5}\right)$ next. We seek $\cos\theta$, where $\sin\theta = \frac{4}{5}$ and θ is in the first quadrant (since the arcsine of a positive number less than 1 is a first quadrant angle). We'll take two approaches:

Method 1: Use trig identities. We have $\cos^2\theta = 1 - \sin^2\theta = 1 - \frac{16}{25} = \frac{9}{25}$. Since θ is in the first quadrant, $\cos^2\theta = \frac{9}{25}$ gives us $\cos\theta = \frac{3}{5}$. So, $\cos\left(\arcsin \frac{4}{5}\right) = \frac{3}{5}$.

Method 2: Use the geometric interpretation of sine and cosine. Suppose $\triangle ABC$ is a right triangle with $\angle C = 90°$ and $\angle A = \theta$, as shown at right. Then, we have $\sin\theta = \frac{BC}{AB}$ and $\cos\theta = \frac{AC}{AB}$. Since we know that $\sin\theta = \frac{4}{5}$, we let $BC = 4$ and $AB = 5$. The Pythagorean Theorem then gives us $AC = 3$, so $\cos\theta = \frac{3}{5}$. (We also could have simply recognized the Pythagorean triple $\{3,4,5\}$.)



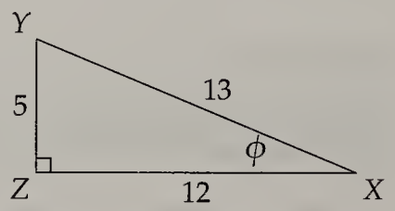
Next, we tackle $\cos\left(\arctan\frac{5}{12}\right)$ and $\sin\left(\arctan\frac{5}{12}\right)$. We let $\phi = \arctan\frac{5}{12}$, so $\tan\phi = \frac{5}{12}$, and $0 < \phi < \frac{\pi}{2}$. Again, we present two solutions:

Method 1: Use trig identities. We could write $\tan\phi$ in terms of $\sin\phi$ and $\cos\phi$, solve for $\sin\phi$ in terms of $\cos\phi$, and then use $\sin^2\phi + \cos^2\phi = 1$. Instead, we'll use $\sec^2\phi = \tan^2\phi + 1$ to find

$$\sec^2\phi = \left(\frac{5}{12}\right)^2 + 1 = \frac{25}{144} + 1 = \frac{169}{144}.$$

Since $0 < \phi < \frac{\pi}{2}$, we know that $\sec\phi > 0$, so we have $\sec\phi = \frac{13}{12}$, which gives us $\cos\phi = \frac{12}{13}$. Finally, we have $\sin^2\phi = 1 - \cos^2\phi = \frac{25}{169}$, so $\sin\phi = \frac{5}{13}$. Therefore, we have $\cos\left(\arctan\frac{5}{12}\right) = \frac{12}{13}$ and $\sin\left(\arctan\frac{5}{12}\right) = \frac{5}{13}$.

Method 2: Geometry. Let $\triangle XYZ$ be a right triangle with $\angle Z = 90°$, $ZY = 5$, $XZ = 12$, and $\angle X = \phi$, as shown at right. We have $\tan\phi = \frac{YZ}{XZ} = \frac{5}{12}$. The Pythagorean Theorem (or knowing the $\{5,12,13\}$ Pythagorean triple) gives us $XY = 13$, so $\sin\phi = \frac{YZ}{XY} = \frac{5}{13}$ and $\cos\phi = \frac{XZ}{XY} = \frac{12}{13}$.

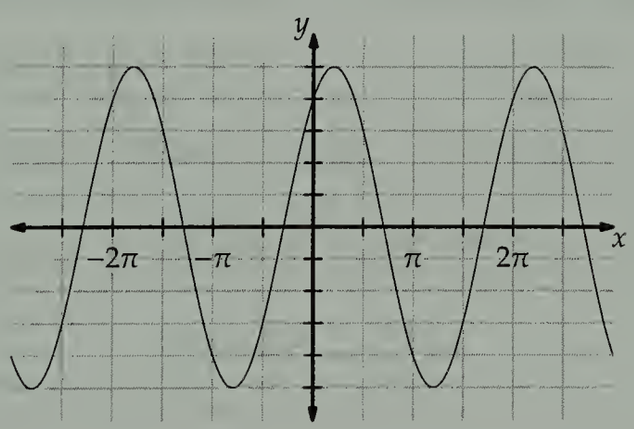


Finally, we have

$$\begin{aligned}\cos\left(\arcsin\frac{4}{5} + \arctan\frac{5}{12}\right) &= \cos\left(\arcsin\frac{4}{5}\right)\cos\left(\arctan\frac{5}{12}\right) - \sin\left(\arcsin\frac{4}{5}\right)\sin\left(\arctan\frac{5}{12}\right)\\ &= \frac{3}{5}\cdot\frac{12}{13} - \frac{4}{5}\cdot\frac{5}{13}\\ &= \frac{16}{65}.\end{aligned}$$

□

Problem 3.13: Below is the graph of $y = 3\sin x + 4\cos x$:



It looks like the graph of a function of the form $a\sin(bx + c)$! Why?

Solution for Problem 3.13: The graph appears to have period 2π and amplitude 5, which might make us first guess that it is the graph of $y = 5\sin x$. The given graph doesn't pass through the origin, so it can't be the graph of $y = 5\sin x$; however, it might be the result of shifting the graph of $y = 5\sin x$ horizontally. If so, then it is the graph

of $y = 5\sin(x+\theta)$ for some angle θ. If this graph is also the graph of $y = 3\sin x + 4\cos x$, then we must have

$$5\sin(x+\theta) = 3\sin x + 4\cos x$$

for all x. Is that possible? Is there any value of θ for which this is true?

It's not clear how we could isolate θ, but we can at least isolate $\sin(x+\theta)$ by dividing both sides by 5, which gives

$$\sin(x+\theta) = \frac{3}{5}\sin x + \frac{4}{5}\cos x.$$

Aha! We saw $\frac{3}{5}$ and $\frac{4}{5}$ in the previous problem. We have $\left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2 = 1$, so we can view $\frac{3}{5}$ and $\frac{4}{5}$ as the cosine and sine of some angle. Using our identity for $\sin(x+\theta)$ tells us just what angle:

$$\sin(x+\theta) = \sin x\cos\theta + \sin\theta\cos x.$$

So, if we let $\theta = \arccos\frac{3}{5}$, which means $\cos\theta = \frac{3}{5}$ and $\sin\theta = \frac{4}{5}$, then we do have $\sin(x+\theta) = \frac{3}{5}\sin x + \frac{4}{5}\cos x$, so

$$5\sin(x+\theta) = 3\sin x + 4\cos x.$$

Mystery solved! □

Problem 3.14: Compute x if $\arctan x + \arctan 1 = 2(\arctan x - \arctan\frac{1}{3})$. *(Source: ARML)*

Solution for Problem 3.14: First, we simplify the equation. We seek x; while it isn't immediately obvious how to isolate x, we can isolate $\arctan x$. Doing so gives us

$$\arctan x = \arctan 1 + 2\arctan\frac{1}{3}.$$

We have $\arctan 1 = \frac{\pi}{4}$, which gives us

$$\arctan x = \frac{\pi}{4} + 2\arctan\frac{1}{3}.$$

But now what?

We'd like to solve for x; now that we've isolated $\arctan x$, we can isolate x by taking the tangent of both sides, which gives us

$$\tan(\arctan x) = \tan\left(\frac{\pi}{4} + 2\arctan\frac{1}{3}\right).$$

By definition, we have $\tan(\arctan x) = x$. We have the tangent of a sum on the right, so we can apply the identity

$$\tan(\alpha+\beta) = \frac{\tan\alpha + \tan\beta}{1 - \tan\alpha\tan\beta}.$$

We then have

$$x = \frac{\tan\frac{\pi}{4} + \tan\left(2\arctan\frac{1}{3}\right)}{1 - \tan\frac{\pi}{4}\tan\left(2\arctan\frac{1}{3}\right)}.$$

Since $\tan\frac{\pi}{4} = 1$, we have

$$x = \frac{1 + \tan\left(2\arctan\frac{1}{3}\right)}{1 - \tan\left(2\arctan\frac{1}{3}\right)}.$$

We're almost there—all we have to do is compute $\tan\left(2\arctan\frac{1}{3}\right)$. Fortunately, the angle sum identity for tangent can do the job again:

$$\begin{aligned}\tan\left(2\arctan\frac{1}{3}\right) &= \tan\left(\arctan\frac{1}{3}+\arctan\frac{1}{3}\right)\\ &= \frac{\tan\left(\arctan\frac{1}{3}\right)+\tan\left(\arctan\frac{1}{3}\right)}{1-\tan\left(\arctan\frac{1}{3}\right)\tan\left(\arctan\frac{1}{3}\right)}\\ &= \frac{\frac{1}{3}+\frac{1}{3}}{1-\frac{1}{3}\cdot\frac{1}{3}}\\ &= \frac{3}{4}.\end{aligned}$$

Now we can find x:

$$x = \frac{1+\tan\left(2\arctan\frac{1}{3}\right)}{1-\tan\left(2\arctan\frac{1}{3}\right)} = \frac{1+\frac{3}{4}}{1-\frac{3}{4}} = 7.$$

□

Exercises

3.2.1 Use the identities we learned in this section to write each of the following in terms of $\cos\theta$, $\sin\theta$, or $\tan\theta$.

(a) $\sin(\pi+\theta)$

(b) $\cos\left(\frac{\pi}{2}+\theta\right)$

(c) $\cos(3\pi-\theta)$

(d) $\tan\left(\theta+\frac{3\pi}{2}\right)$

3.2.2 Show that $\tan(\alpha-\beta) = \dfrac{\tan\alpha-\tan\beta}{1+\tan\alpha\tan\beta}$.

3.2.3 Evaluate $\tan 285°$.

3.2.4 Prove that $\sin(\alpha-\beta)\sin(\alpha+\beta) = \sin^2\alpha-\sin^2\beta$.

3.2.5 Find all acute angles x such that $\sin 2x\sin 3x = \cos 2x\cos 3x$. *(Source: AHSME)*

3.2.6 Angles A and B are acute angles such that $\tan A+\tan B+\tan A\tan B = 1$. If $A-B = 41°$, compute the measure of angle A. *(Source: NYSML)*

3.2.7 Suppose that $\sin a+\sin b = \sqrt{5/3}$ and $\cos a+\cos b = 1$. What is $\cos(a-b)$? *(Source: AMC)*

3.2.8★ In the text, we proved that $\sin(\alpha+\beta) = \sin\alpha\cos\beta+\sin\beta\cos\alpha$ only for the case where α and β are acute angles such that $\alpha+\beta$ is also acute. In this problem, we explore the type of careful casework we can use to prove that this identity holds for all angles. We will prove that the identity holds if α and β are acute but $\alpha+\beta$ is obtuse.

(a) Let $\gamma = \frac{\pi}{2}-\alpha$ and $\phi = \frac{\pi}{2}-\beta$. Why must γ and ϕ be acute?

(b) Write $\sin(\alpha+\beta)$ in terms of γ, ϕ, and π.

(c) Show that $\sin(\alpha+\beta) = \sin(\gamma+\phi)$.

(d) Express $\sin(\gamma+\phi)$ in terms of $\sin\alpha$, $\sin\beta$, $\cos\alpha$, and $\cos\beta$ to complete the proof of the identity for the case in which α and β are acute but $\alpha+\beta$ is obtuse. Why is it important in this step that γ and ϕ are acute?

(e)★ Show that the identity holds if α and β are both obtuse.

3.3 Double and Half Angles

In Problem 3.14, we used the identity for the tangent of a sum of angles to evaluate $\tan\left(2\arctan\frac{1}{3}\right)$. In other words, we used the angle sum identity to evaluate the tangent of double an angle. Double angles and half angles occur frequently in trigonometry; in this section we develop and apply identities for trig functions of both.

Problems

Problem 3.15: Find $\sin 2x$ and $\cos 2x$ in terms of $\sin x$ and $\cos x$. Test your identities by finding $\sin 120°$ and $\cos 120°$ with them.

Problem 3.16: Find $\tan 2x$ in terms of $\tan x$.

Problem 3.17: Use your identity for $\cos 2x$ in Problem 3.15 together with $\sin^2 x + \cos^2 x = 1$ to find $\cos(x/2)$ and $\sin(x/2)$ in terms of $\cos x$. Test your identities by finding $\cos 120°$ and $\sin 120°$ by letting $x = 240°$ in your identities.

Problem 3.18: Find $\cos 15°$ using an identity you found in the previous problem. Compare your answer to the result in Problem 3.11.

Problem 3.19: Show that $\tan\dfrac{x}{2} = \dfrac{\sin x}{1+\cos x} = \dfrac{1-\cos x}{\sin x}$.

Problem 3.20: Compute the number of degrees in the smallest positive angle x that satisfies the equation $8\sin x\cos^5 x - 8\sin^5 x\cos x = 1$. *(Source: ARML)*

Problem 3.21: Find $(\cos 20°)(\cos 40°)(\cos 80°)$.

Problem 3.15: Find $\sin 2x$ and $\cos 2x$ in terms of $\sin x$ and $\cos x$. Test your identities by finding $\sin 120°$ and $\cos 120°$ with them.

Solution for Problem 3.15: Applying the angle sum identities for sine and cosine, we have

$$\sin 2x = \sin(x+x) = \sin x\cos x + \sin x\cos x = 2\sin x\cos x,$$
$$\cos 2x = \cos(x+x) = \cos x\cos x - \sin x\sin x = \cos^2 x - \sin^2 x.$$

We can use the identity $\sin^2 x + \cos^2 x = 1$ to write $\cos 2x$ solely in terms of $\cos x$ or in terms of $\sin x$. Substituting $\cos^2 x = 1 - \sin^2 x$ gives

$$\cos 2x = \cos^2 x - \sin^2 x = (1-\sin^2 x) - \sin^2 x = 1 - 2\sin^2 x,$$

and substituting $\sin^2 x = 1 - \cos^2 x$ gives

$$\cos 2x = \cos^2 x - \sin^2 x = \cos^2 x - (1-\cos^2 x) = 2\cos^2 x - 1.$$

We can test these by letting $x = 60°$. Then, we have $\sin 2x = \sin 120° = \frac{\sqrt{3}}{2}$ and $2\sin x\cos x = 2\cdot\frac{\sqrt{3}}{2}\cdot\frac{1}{2} = \frac{\sqrt{3}}{2}$. We also have $\cos 2x = \cos 120° = -\frac{1}{2}$ and $\cos^2 x - \sin^2 x = \frac{1}{4} - \frac{3}{4} = -\frac{1}{2}$. So, both our identities work for $x = 120°$. □

Problem 3.16: Find $\tan 2x$ in terms of $\tan x$.

Solution for Problem 3.16: Applying the angle sum identity for tangent, we have

$$\tan 2x = \tan(x+x) = \frac{\tan x + \tan x}{1 - \tan x \tan x} = \frac{2\tan x}{1 - \tan^2 x}.$$

□

Important: **Double angle identities:**

$$\sin 2x = 2\sin x \cos x,$$
$$\cos 2x = \cos^2 x - \sin^2 x = 2\cos^2 x - 1 = 1 - 2\sin^2 x,$$
$$\tan 2x = \frac{2\tan x}{1 - \tan^2 x}.$$

You'll use these identities enough that they're worth memorizing. (In fact, you'll likely use them so much in this chapter that you'll have them memorized by the end of the chapter.)

Problem 3.17: Find $\cos\frac{x}{2}$ and $\sin\frac{x}{2}$ in terms of $\cos x$.

Solution for Problem 3.17: We don't yet have any identities involving half angles. However, we do have double angle identities, which means we can express a trig function of an angle in terms of trig functions of half that angle. For example, if we start with

$$\sin 2\theta = 2\sin\theta\cos\theta,$$

and let $\theta = \frac{x}{2}$, we have

$$\sin x = 2\sin\frac{x}{2}\cos\frac{x}{2}.$$

We now have an angle of $\frac{x}{2}$, but we unfortunately have both the cosine and sine of this angle. This happened because our identity for $\sin 2\theta$ includes both $\sin\theta$ and $\cos\theta$. We'd prefer to use a double angle identity that only has one of $\sin\theta$ or $\cos\theta$, not both. Fortunately, $\cos 2\theta$ offers such identities:

$$\cos 2\theta = 2\cos^2\theta - 1 = 1 - 2\sin^2\theta.$$

Letting $\theta = \frac{x}{2}$ in $\cos 2\theta = 2\cos^2\theta - 1$ gives

$$\cos x = 2\cos^2\frac{x}{2} - 1.$$

Solving for $\cos\frac{x}{2}$ gives

$$\cos\frac{x}{2} = \pm\sqrt{\frac{1+\cos x}{2}}.$$

The $\pm$ here does not mean that $\cos\frac{x}{2}$ has two values! It means that we will take either the positive or the negative square root, depending on what quadrant $\frac{x}{2}$ is in. For example, if $x = 240°$, then $\frac{x}{2} = 120°$, which is a second quadrant angle. Therefore, $\cos 120°$ is negative, and we must take the negative root in the identity. Letting $x = 240°$ in the identity gives

$$\cos 120° = -\sqrt{\frac{1+\cos 240°}{2}} = -\sqrt{\frac{1-\frac{1}{2}}{2}} = -\sqrt{\frac{1}{4}} = -\frac{1}{2},$$

as expected.

We deal with $\sin\frac{x}{2}$ by letting $\theta = \frac{x}{2}$ in $\cos 2\theta = 1 - 2\sin^2\theta$, which gives us

$$\cos x = 1 - 2\sin^2\frac{x}{2}.$$

Solving for $\sin\frac{x}{2}$ gives

$$\sin\frac{x}{2} = \pm\sqrt{\frac{1-\cos x}{2}}.$$

Once again, the $\pm$ is a reminder that we must take the positive or negative root depending on the quadrant of $\frac{x}{2}$. We again consider the example $x = 240°$, which makes $\frac{x}{2} = 120°$, a second quadrant angle for which sine is positive. Therefore, we have

$$\sin 120° = \sqrt{\frac{1-\cos 120°}{2}} = \sqrt{\frac{1-\left(-\frac{1}{2}\right)}{2}} = \sqrt{\frac{3}{4}} = \frac{\sqrt{3}}{2},$$

as expected. □

Let's try putting our half angle identities to work on an angle we can't tackle with the unit circle.

Problem 3.18: Find $\cos 15°$ using an identity you found in the previous problem. Compare your answer to the result in Problem 3.11.

Solution for Problem 3.18: Earlier, we used the angle difference identities to find $\cos 15° = \cos(45° - 30°) = \frac{\sqrt{6}+\sqrt{2}}{4}$. We now can use half angle identities to compute $\cos 15°$ as $\cos\frac{30°}{2}$. Since $15°$ is a first quadrant angle, we take the positive root, and we have

$$\cos 15° = \cos\frac{30°}{2} = \sqrt{\frac{1+\cos 30°}{2}} = \sqrt{\frac{1+\frac{\sqrt{3}}{2}}{2}} = \sqrt{\frac{2+\sqrt{3}}{4}} = \frac{\sqrt{2+\sqrt{3}}}{2}.$$

Uh-oh! This doesn't look anything like the result we found with $\cos(45° - 30°)$. Did we do something wrong?

If we calculate both $\frac{\sqrt{6}+\sqrt{2}}{4}$ and $\frac{\sqrt{2+\sqrt{3}}}{2}$ with a calculator, it appears that the two are equal. They sure don't look equal! To confirm by hand that they are equal, we get rid of some of the square roots by squaring both numbers:

$$\left(\frac{\sqrt{6}+\sqrt{2}}{4}\right)^2 = \frac{6+2\sqrt{6}\sqrt{2}+2}{16} = \frac{8+2\cdot 2\sqrt{3}}{16} = \frac{2+\sqrt{3}}{4},$$

$$\left(\frac{\sqrt{2+\sqrt{3}}}{2}\right)^2 = \frac{2+\sqrt{3}}{4}.$$

Now we see the whole story. Our two results for $\cos 15°$ are positive and have equal squares, so the two must be the same. □

Problem 3.19: Show that $\tan\frac{x}{2} = \frac{\sin x}{1+\cos x} = \frac{1-\cos x}{\sin x}$.

Solution for Problem 3.19: We already have $\cos\frac{x}{2}$ and $\sin\frac{x}{2}$, so we can find $\tan\frac{x}{2}$:

$$\tan\frac{x}{2} = \frac{\sin\frac{x}{2}}{\cos\frac{x}{2}} = \frac{\pm\sqrt{\frac{1-\cos x}{2}}}{\pm\sqrt{\frac{1+\cos x}{2}}} = \pm\frac{\sqrt{1-\cos x}}{\sqrt{1+\cos x}}.$$

Again, we include $\pm$ to indicate that we we'll use the quadrant of $\frac{x}{2}$ to determine whether to take the positive or the negative result. We can simplify our expression a bit by getting rid of the square root in the denominator. We do so by multiplying the numerator and denominator by $\sqrt{1+\cos x}$. This gives us

$$\begin{aligned}\tan\frac{x}{2} &= \pm\frac{\sqrt{1-\cos x}}{\sqrt{1+\cos x}}\cdot\frac{\sqrt{1+\cos x}}{\sqrt{1+\cos x}}\\ &= \pm\frac{\sqrt{(1-\cos x)(1+\cos x)}}{1+\cos x}\\ &= \pm\frac{\sqrt{1-\cos^2 x}}{1+\cos x}.\end{aligned}$$

We recognize $1-\cos^2 x$ as equal to $\sin^2 x$, so we can simplify even more:

$$\tan\frac{x}{2} = \pm\frac{\sqrt{1-\cos^2 x}}{1+\cos x} = \pm\frac{\sqrt{\sin^2 x}}{1+\cos x} = \pm\frac{\sin x}{1+\cos x}.$$

We should worry about what happens if the denominator on the right side is 0. If $1+\cos x = 0$, then $\cos x = -1$, which means x is an odd integer multiple of π. For these values of x, $\tan\frac{x}{2}$ is indeed undefined, because $\cos\frac{x}{2} = 0$ whenever x is an odd integer multiple of π.

We might think we're finished, but then we notice that the problem claims that

$$\tan\frac{x}{2} = \frac{\sin x}{1+\cos x}.$$

Our result above has a $\pm$ in front of the right side. Is it true that we can drop the $\pm$? The value of $1+\cos x$ is always nonnegative, since $\cos x \geq -1$. So, to compare the sign of $\tan\frac{x}{2}$ to the sign of $\frac{\sin x}{1+\cos x}$, we need only compare the signs of $\tan\frac{x}{2}$ and $\sin x$. We investigate by considering all possible values of x.

If $\sin x = 1$ or if x is in the first or second quadrant, then $x = \theta + 2k\pi$ for some integer k and some θ with $0 < \theta < \pi$. Then, $\sin x$ is positive, and $\frac{x}{2} = \frac{\theta}{2} + k\pi$. Since $0 < \frac{\theta}{2} < \frac{\pi}{2}$, the angle $\frac{\theta}{2} + k\pi$ is in the first or third quadrant, which means $\tan\frac{x}{2}$ is also positive.

If $\sin x = -1$ or if x is in the third or fourth quadrant, then $x = \theta + 2k\pi$ for some integer k and some θ with $\pi < \theta < 2\pi$. This means $\sin x$ is negative and $\frac{x}{2} = \frac{\theta}{2} + k\pi$. Since $\frac{\pi}{2} < \frac{\theta}{2} < \pi$, the angle $\frac{\theta}{2} + k\pi$ is in the second or fourth quadrant, which means $\tan\frac{x}{2}$ is also negative.

Finally, we note that if x is an even multiple of π, then $\tan\frac{x}{2} = 0$ and $\sin x = 0$. (We noted above that if x is an odd multiple of π, then both $\tan\frac{x}{2}$ and $(\sin x)/(1+\cos x)$ are undefined.)

The sign of $\tan\frac{x}{2}$ therefore matches that of $\sin x$ for all possible values of x for which $\tan\frac{x}{2}$ is defined, so we have

$$\tan\frac{x}{2} = \frac{\sin x}{1+\cos x}.$$

We multiply the numerator and denominator of the right side by $1-\cos x$ to produce the other desired identity:

$$\tan\frac{x}{2} = \frac{\sin x}{1+\cos x}\cdot\frac{1-\cos x}{1-\cos x} = \frac{(\sin x)(1-\cos x)}{1-\cos^2 x} = \frac{(\sin x)(1-\cos x)}{\sin^2 x} = \frac{1-\cos x}{\sin x}.$$

□

Extra! Challenge: Prove that

$$\arctan 1 + \arctan 2 + \arctan 3 = \pi$$

by drawing three well-chosen triangles on the Cartesian plane. *Answer on page 110.*

Important: **Half angle identities**:

$$\cos\frac{\theta}{2} = \pm\sqrt{\frac{1+\cos\theta}{2}},$$
$$\sin\frac{\theta}{2} = \pm\sqrt{\frac{1-\cos\theta}{2}},$$
$$\tan\frac{\theta}{2} = \frac{\sin\theta}{1+\cos\theta} = \frac{1-\cos\theta}{\sin\theta}.$$

The $\pm$ symbols on the first two are reminders that we must use the quadrant of $\frac{\theta}{2}$ to determine the appropriate sign.

If you know the double angle identities, then you should be able to quickly derive these half angle identities whenever you need them.

Problem 3.20: Compute the number of degrees in the smallest positive angle x that satisfies the equation $8\sin x\cos^5 x - 8\sin^5 x\cos x = 1$. *(Source: ARML)*

Solution for Problem 3.20: Ultimately, our goal is to isolate a trig function, but the expression on the left is pretty complicated. Most intimidating are the fifth powers. We start chipping away at these in the same way we deal with high degree polynomials, by factoring. First, we factor out $8\sin x\cos x$, and we have

$$8\sin x\cos x(\cos^4 x - \sin^4 x) = 1.$$

Now, we have a difference of squares: we can factor $\cos^4 x - \sin^4 x$ as

$$\cos^4 x - \sin^4 x = (\cos^2 x)^2 - (\sin^2 x)^2 = (\cos^2 x - \sin^2 x)(\cos^2 x + \sin^2 x).$$

Moreover, $\cos^2 x + \sin^2 x = 1$, so we have $\cos^4 x - \sin^4 x = \cos^2 x - \sin^2 x$, which means

$$8\sin x\cos x(\cos^4 x - \sin^4 x) = 8\sin x\cos x(\cos^2 x - \sin^2 x).$$

Our equation now is

$$8\sin x\cos x(\cos^2 x - \sin^2 x) = 1.$$

We've at least reduced the powers of the trig functions, but we still have some work to do. We'd like to simplify the equation until we have just a single trig function set equal to a constant. We recognize some expressions on the left side that might help us continue simplifying. Specifically, we recognize $\sin x\cos x$ from the double angle identity for sine, and $\cos^2 x - \sin^2 x$ from the double angle identity for cosine. Applying these identities "in reverse" gives us $\sin x\cos x = \frac{1}{2}\sin 2x$ and $\cos^2 x - \sin^2 x = \cos 2x$. Now, our equation is $8\left(\frac{1}{2}\sin 2x\right)\cos 2x = 1$, or

$$4\sin 2x\cos 2x = 1.$$

Once again, we have the product of sine and cosine of the same angle, which makes us think of the double angle identity for sine. We have $\sin 2x\cos 2x = \frac{1}{2}\sin 2(2x) = \frac{1}{2}\sin 4x$, so now our equation is $4\left(\frac{1}{2}\sin 4x\right) = 1$, or $2\sin 4x = 1$. Therefore, we have $\sin 4x = \frac{1}{2}$. We can use what we know about sine to finish. (This is why we wanted to manipulate the equation into a single trig function set equal to a constant.) The smallest positive angle with sine equal to $\frac{1}{2}$ is $30°$, so we have $4x = 30°$, from which we find $x = \frac{15}{2}°$. □

The key step in our solution is recognizing that expressions in the problem match the forms of the double angle identities.

Concept: Recognizing the forms of common trigonometric identities can help solve many trigonometric problems.

Problem 3.21: Find $(\cos 20°)(\cos 40°)(\cos 80°)$.

Solution for Problem 3.21: This problem doesn't match the form of any of our identities, but the facts that $80° = 2\cdot 40°$ and $40° = 2\cdot 20°$ make us try applying the double angle identities. We have cosines, so let's try applying the cosine double angle identity to the latter two cosines:

$$(\cos 20°)(\cos 40°)(\cos 80°) = (\cos 20°)(2\cos^2 20° - 1)(2\cos^2 40° - 1).$$

Yuck. And if we apply this again to the $\cos 40°$ on the right, it will get even uglier.

Back to the drawing board. Those angles really make us want to use the double angle identities. The cosine double angle identity was a loser, but there is a cosine in the sine double angle identity. For example, we have

$$\sin 40° = 2\cos 20°\sin 20°,$$

so $\cos 20° = \dfrac{\sin 40°}{2\sin 20°}$. Making this substitution gives us

$$(\cos 20°)(\cos 40°)(\cos 80°) = \left(\frac{\sin 40°}{2\sin 20°}\right)(\cos 40°)(\cos 80°) = \frac{(\sin 40°)(\cos 40°)(\cos 80°)}{2\sin 20°}.$$

At first, this doesn't look helpful, but then we see the product of sine and cosine of the same angle, which means we can use the sine double angle identity again. Running this identity "in reverse," as in Problem 3.20, we have $(\sin 40°)(\cos 40°) = \frac{\sin 80°}{2}$, so we have

$$(\cos 20°)(\cos 40°)(\cos 80°) = \frac{(\sin 40°)(\cos 40°)(\cos 80°)}{2\sin 20°} = \frac{(\sin 80°)(\cos 80°)}{4\sin 20°}.$$

We can do it again! Since $(\sin 80°)(\cos 80°) = \frac{1}{2}\sin 160°$, we have

$$(\cos 20°)(\cos 40°)(\cos 80°) = \frac{(\sin 80°)(\cos 80°)}{4\sin 20°} = \frac{\sin 160°}{8\sin 20°}.$$

Now we need to relate $\sin 160°$ and $\sin 20°$. We got this far by recognizing relationships among angles in the problem. We do the same here. We note that $160° + 20° = 180°$. So, we apply the identity $\sin(180° - x) = \sin x$ to see that $\sin 160° = \sin 20°$, which means these two cancel in the fraction above and leave the surprising result

$$(\cos 20°)(\cos 40°)(\cos 80°) = \frac{1}{8}.$$

□

Concept: Recognizing relationships among angles in a complicated trig expression can help simplify the expression.

Exercises

3.3.1 Evaluate $\sin 22.5°$.

3.3.2 If $\sin\theta = \frac{1}{4}$, then what is $\cos 2\theta$?

3.3.3 Prove that $\sin 2x = \dfrac{2\tan x}{1+\tan^2 x}$ for all x for which $\tan x$ is defined.

3.3.4 Show that $\tan x + \cot x = 2\csc 2x$.

3.3.5 If $-\pi/2 < A < \pi/2$ and $A \neq -\pi/4$, then prove that $\dfrac{\sin 2A - \cos 2A + 1}{\sin 2A + \cos 2A + 1} = \tan A$. *(Source: CEMC)*

3.3.6 If $\sin 2\theta = 21/25$ and $\cos\theta > \sin\theta$, then compute $\cos\theta - \sin\theta$. *(Source: Mandelbrot)*

3.3.7 Solve the equation $\cos\theta + \sin\theta = \sqrt{3/2}$ for $0 \leq \theta < 2\pi$.

3.3.8 In this problem, we prove the **triple angle identities** for sine and cosine.

(a) Prove that $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$ for all angles θ.

(b) Find a similar identity for $\sin 3\theta$ by expressing $\sin 3\theta$ in terms of $\sin\theta$.

(c)★ Suppose $\dfrac{\cos 3x}{\cos x} = \dfrac{1}{3}$ for some angle x, with $0 \leq x \leq \frac{\pi}{2}$. Determine $\dfrac{\sin 3x}{\sin x}$. *(Source: USAMTS)*

3.4 Sum-to-product and Product-to-sum

Problems

Problem 3.22:
(a) Express $\cos\alpha\cos\beta$ in terms of $\cos(\alpha+\beta)$ and $\cos(\alpha-\beta)$.
(b) Find a similar identity for $\sin\alpha\sin\beta$.
(c) Find a similar identity for $\sin\alpha\cos\beta$.

Problem 3.23: In Problem 3.21, we proved that $(\cos 20°)(\cos 40°)(\cos 80°) = \frac{1}{8}$ using double angle identities. Use one of the identities from Problem 3.22 to find another proof.

Problem 3.24:
(a) Use an identity from Problem 3.22 to show that $\cos x + \cos y = 2\cos\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right)$.
(b) Find a similar identity to that in part (a) for $\cos x - \cos y$.
(c) Find identities for $\sin x + \sin y$ and $\sin x - \sin y$.

Problem 3.25: Compute $\dfrac{\sin 13° + \sin 47° + \sin 73° + \sin 107°}{\cos 17°}$. *(Source: ARML)*

Problem 3.22:
(a) Express $\cos\alpha\cos\beta$ in terms of $\cos(\alpha+\beta)$ and $\cos(\alpha-\beta)$.
(b) Find a similar identity for $\sin\alpha\sin\beta$.
(c) Find a similar identity for $\sin\alpha\cos\beta$.

Solution for Problem 3.22:

(a) We first write out the angle sum and angle difference identities for cosine:

$$\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta,$$
$$\cos(\alpha-\beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta.$$

Adding them cancels $\sin\alpha\sin\beta$ and leaves us with $\cos(\alpha+\beta)+\cos(\alpha-\beta) = 2\cos\alpha\cos\beta$. Dividing by 2 gives

$$\cos\alpha\cos\beta = \frac{1}{2}\left(\cos(\alpha+\beta)+\cos(\alpha-\beta)\right).$$

(b) If we subtract $\cos(\alpha-\beta)$ from $\cos(\alpha+\beta)$ instead of adding, then we cancel $\cos\alpha\cos\beta$, and leave

$$\cos(\alpha+\beta)-\cos(\alpha-\beta) = -2\sin\alpha\sin\beta.$$

Dividing by -2 gives

$$\sin\alpha\sin\beta = \frac{1}{2}\left(\cos(\alpha-\beta)-\cos(\alpha+\beta)\right).$$

(c) We don't have a product of sine and cosine in our angle sum or difference identities for cosine, but we do have them in the corresponding sine identities:

$$\sin(\alpha+\beta) = \sin\alpha\cos\beta+\sin\beta\cos\alpha,$$
$$\sin(\alpha-\beta) = \sin\alpha\cos\beta-\sin\beta\cos\alpha.$$

Adding these and dividing by 2 gives us

$$\sin\alpha\cos\beta = \frac{1}{2}\left(\sin(\alpha+\beta)+\sin(\alpha-\beta)\right).$$

□

We call the identities we found in Problem 3.22 **product-to-sum** and **product-to-difference identities**.

Important: **Product-to-sum and product-to-difference identities**:

$$\cos\alpha\cos\beta = \frac{1}{2}\left(\cos(\alpha+\beta)+\cos(\alpha-\beta)\right),$$
$$\sin\alpha\sin\beta = \frac{1}{2}\left(\cos(\alpha-\beta)-\cos(\alpha+\beta)\right),$$
$$\sin\alpha\cos\beta = \frac{1}{2}\left(\sin(\alpha+\beta)+\sin(\alpha-\beta)\right).$$

Again, there's nothing to memorize here: these are a quick, direct result of the angle sum and angle difference identities. The main idea here is that we can turn a product of cosines or sines into a sum or difference. Sometimes this can be particularly convenient, especially when applying the identity gives us angles we know how to handle.

Problem 3.23: In Problem 3.21, we proved that $(\cos 20°)(\cos 40°)(\cos 80°) = \frac{1}{8}$ using double angle identities. Use one of the identities from Problem 3.22 to find another proof.

Solution for Problem 3.23: We now have an identity for the product of two cosines, but in this problem, we have the product of three cosines. To which two should we apply the identity?

Concept: If a problem involves the product of sines and/or cosines of angles with a convenient sum or difference, consider the product-to-sum identities.

We apply the identity to $(\cos 40°)(\cos 80°)$, since the angles in these have sum $120°$, which we know to handle inside trig functions. (See if you can find another solution by applying a product-to-sum identity to $(\cos 20°)(\cos 40°)$.) We have

$$(\cos 80°)(\cos 40°) = \frac{1}{2}\left(\cos(80°-40°)+\cos(80°+40°)\right) = \frac{1}{2}\left(\cos 40°+\cos 120°\right) = \frac{1}{2}\left(\cos 40°-\frac{1}{2}\right),$$

so now we have

$$(\cos 20°)(\cos 40°)(\cos 80°) = \frac{1}{2}(\cos 20°)\left(\cos 40° - \frac{1}{2}\right).$$

If we expand the product on the right, we'll have another product of cosines with a convenient sum.

Concept: If a tactic makes some progress on a problem, but doesn't entirely solve it, look for a way to use the tactic again.

Expanding the product gives us

$$(\cos 20°)(\cos 40°)(\cos 80°) = \frac{1}{2}(\cos 20°)\left(\cos 40° - \frac{1}{2}\right) = \frac{1}{2}\left((\cos 40°)(\cos 20°) - \frac{1}{2}\cos 20°\right).$$

Applying a product-to-sum identity to $(\cos 40°)(\cos 20°)$ gives

$$(\cos 40°)(\cos 20°) = \frac{1}{2}\left(\cos(40° + 20°) + \cos(40° - 20°)\right) = \frac{1}{2}\left(\frac{1}{2} + \cos 20°\right),$$

so we now have

$$\begin{aligned}(\cos 20°)(\cos 40°)(\cos 80°) &= \frac{1}{2}\left((\cos 40°)(\cos 20°) - \frac{1}{2}\cos 20°\right)\\ &= \frac{1}{2}\left(\frac{1}{2}\left(\frac{1}{2} + \cos 20°\right) - \frac{1}{2}\cos 20°\right)\\ &= \frac{1}{2}\left(\frac{1}{4} + \frac{1}{2}\cos 20° - \frac{1}{2}\cos 20°\right)\\ &= \frac{1}{8}.\end{aligned}$$

□

The two very different-looking proofs of $(\cos 20°)(\cos 40°)(\cos 80°) = \frac{1}{8}$ exemplify how deeply related trigonometric identities are to each other. Many trig problems can be solved in a variety of ways using different combinations of identities, so don't get locked into looking for *the* solution—there may be several approaches that work.

Problem 3.24:

(a) Use an identity from Problem 3.22 to show that $\cos x + \cos y = 2\cos\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right)$.

(b) Find an identity similar to that in part (a) for $\cos x - \cos y$.

(c) Find identities for $\sin x + \sin y$ and $\sin x - \sin y$.

Solution for Problem 3.24:

(a) We already have an identity that relates a product of cosines to a sum of cosines: our product-to-sum identity. We'll write it with the sum isolated on the left, since that's the form of the desired identity in this problem:

$$\cos(\alpha + \beta) + \cos(\alpha - \beta) = 2\cos\alpha\cos\beta.$$

To convert this to an identity for $\cos x + \cos y$, we simply let $\alpha + \beta = x$ and $\alpha - \beta = y$. Adding these gives $2\alpha = x + y$, so $\alpha = (x + y)/2$. Subtracting $\alpha - \beta = y$ from $\alpha + \beta = x$ gives $\beta = (x - y)/2$. Substituting these expressions for α and β in the identity above gives the desired

$$\cos x + \cos y = 2\cos\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right).$$

(b) The previous part gives us a clear path to find an identity for $\cos x - \cos y$. We start with

$$\cos(\alpha - \beta) - \cos(\alpha + \beta) = 2\sin\alpha\sin\beta.$$

Substituting $x = \alpha + \beta$ and $y = \alpha - \beta$ in this identity, so that $\alpha = (x+y)/2$ and $\beta = (x-y)/2$ as before, gives us

$$\cos y - \cos x = 2\sin\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right),$$

or

$$\cos x - \cos y = -2\sin\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right).$$

(c) For these, we turn to the product-to-sum identity that features a sum of sines. As before, we isolate the sum on the left:

$$\sin(\alpha + \beta) + \sin(\alpha - \beta) = 2\sin\alpha\cos\beta.$$

Once again, we use the substitution $x = \alpha + \beta$ and $y = \alpha - \beta$, so $\alpha = (x+y)/2$ and $\beta = (x-y)/2$, and we have

$$\sin x + \sin y = 2\sin\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right).$$

We don't have an identity with a difference of sines among those we found in Problem 3.22, but we can convert the sine sum above to a sine difference by noting that $\sin(-z) = -\sin z$. So, letting $y = -z$ in the sine sum identity above gives us

$$\sin x - \sin z = 2\sin\left(\frac{x-z}{2}\right)\cos\left(\frac{x+z}{2}\right).$$

□

Important: **Sum-to-product and difference-to-product identities:**

$$\cos x + \cos y = 2\cos\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right),$$
$$\cos x - \cos y = -2\sin\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right),$$
$$\sin x + \sin y = 2\sin\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right),$$
$$\sin x - \sin y = 2\sin\left(\frac{x-y}{2}\right)\cos\left(\frac{x+y}{2}\right).$$

As with many of the other identities we've studied, you don't need to memorize these. We've seen that these are a direct result of the product-to-sum and product-to-difference identities, which are in turn a straightforward manipulation of the angle sum and angle difference identities.

Problem 3.25: Compute $\dfrac{\sin 13^\circ + \sin 47^\circ + \sin 73^\circ + \sin 107^\circ}{\cos 17^\circ}$. *(Source: ARML)*

Solution for Problem 3.25: This is a natural candidate for sum-to-product identities, both because we have a sum of sines in the numerator, and because the angles have convenient sums. But how should we pair up the sines to apply the sum of sines identity? We could pair the first two and the last two, since $13^\circ + 47^\circ = 60^\circ$ and $73^\circ + 107^\circ = 180^\circ$, or we could pair the outer two and the inner two, since $13^\circ + 107^\circ = 47^\circ + 73^\circ = 120^\circ$. Looking at the differences of the paired angles gives us the answer. In our first option, these differences are both 34°, while in the second option, we have $107 - 13^\circ = 94^\circ$ and $73^\circ - 47^\circ = 26^\circ$. While 34°, 94°, and 26° are all inconvenient, we do note that 34° is double the 17° in the denominator.

Concept: In problems involving several trig functions of inconvenient angles, look for relationships among the angles.

Applying the sine sum-to-product identities to $\sin 13° + \sin 47°$ and $\sin 73° + \sin 107°$, we have

$$\begin{aligned}\frac{\sin 13° + \sin 47° + \sin 73° + \sin 107°}{\cos 17°} &= \frac{2\sin 30° \cos 17° + 2\sin 90° \cos 17°}{\cos 17°} \\ &= \frac{\cos 17° \left(2 \cdot \frac{1}{2} + 2 \cdot 1\right)}{\cos 17°} \\ &= 3.\end{aligned}$$

□

Exercises

3.4.1 Show that $(\cos 4x - \cos 2x)/(2\sin 3x) = -\sin x$ for all x such that $\sin 3x \neq 0$.

3.4.2 Find the acute angle θ such that $\cos 27° + \cos 33° = \sqrt{3}\cos\theta$.

3.4.3 Consider the following two identities:

$$\sin(\alpha + \beta) = \sin\alpha\cos\beta + \sin\beta\cos\alpha,$$
$$\sin(\alpha - \beta) = \sin\alpha\cos\beta - \sin\beta\cos\alpha.$$

If we subtract the second equation from the first, we get an identity for $\sin\beta\cos\alpha$. But we already have an identity for $\sin\alpha\cos\beta$, and it doesn't look just like the result we get for $\sin\beta\cos\alpha$ when we subtract above. Explain.

3.4.4 Express

$$\frac{\sin 10° + \sin 20° + \sin 30° + \sin 40° + \sin 50° + \sin 60° + \sin 70° + \sin 80°}{\cos 5° \cos 10° \cos 20°}$$

without using trigonometric functions. *(Source: HMMT)*

3.4.5 Prove that

$$4\sin A \sin B \sin C = -\sin(A + B + C) + \sin(B + C - A) + \sin(C + A - B) + \sin(A + B - C)$$

for all angles A, B, and C.

3.4.6⋆ Find the smallest positive integer solution to

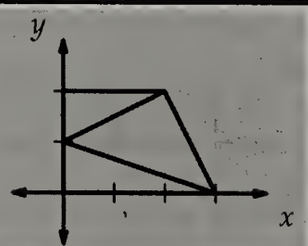
$$\tan 19x° = \frac{\cos 96° + \sin 96°}{\cos 96° - \sin 96°}.$$

(Source: AIME) **Hints:** 282, 24

Extra! The diagram at right can be used to prove that

$$\arctan 1 + \arctan 2 + \arctan 3 = \pi.$$

Don't see how? Here's a hint: All three triangles are right triangles, and all three have an acute angle with a vertex at $(0, 1)$. What is the tangent of each of these acute angles?



Sidenote:

Loosely speaking, sound is caused by vibrations that are sinusoidal in nature; these cause what are called **sound waves**. These waves have an amplitude (which is related to the volume of the sound) and a frequency. The frequency of a sound is related to the **pitch** of the sound, which is how high or low the sound is. As we noted on page 61, the frequency of a sinusoidal function of time is the reciprocal of the function's period. So, the period of a sound wave is related to the pitch of the sound.

If you've ever tuned a guitar by striking two strings and listening to how the sounds they make interact, you've heard the sum-to-product rule in action. If the guitar is in tune, then you hear a nice, even sound. But if you have the guitar very nearly in tune, but not quite, then the sound oscillates. If the guitar is further out of tune, then the sound oscillates a little more. Let's see if we can explain what's going here.

Each string produces a sound wave that is of the form $f(t) = \sin at$ for some constant a. If the guitar is perfectly in tune, then the two waves don't interfere at all, and you hear a nice, smooth sound. We can get an example of what happens here by looking at the graph of $y = \sin 30x + \sin 30x$:

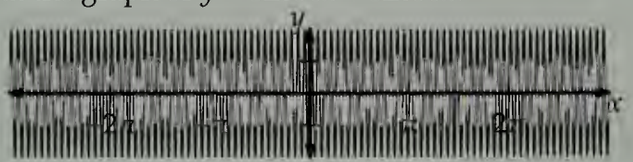


Figure 3.6: Graph of $y = \sin 30x + \sin 30x$

That's a nice, smooth, even sound—the amplitude of the wave is constant throughout. But what happens if they are a little off? Below is the graph of $y = \sin 31x + \sin 29x$:

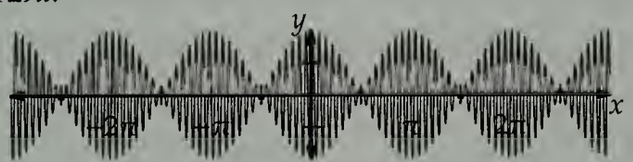


Figure 3.7: Graph of $y = \sin 31x + \sin 29x$

Here, we can see what you hear when the strings are just a little bit out of tune. Instead of a nice, even oscillation, there is an overall up and down to the intensity of the sound. When the strings are a little more out of tune, this up and down happens more quickly, so the sound oscillates faster. Let's graph $y = \sin 32x + \sin 28x$ to get a picture of what's happening:

Figure 3.8: Graph of $y = \sin 32x + \sin 28x$

See if you can use the sum-to-product identity to figure out what's going on here.

Answer on page 127

3.5⋆ Problem Solving with Identities

We've solved a number of challenging problems with trig identities. In each case, the section in which the problem appeared gave us a very convenient hint which identity to try. In this section, we tackle problems without this helping hand, and in doing so, learn some strategies for determining which identity will be helpful on a given problem.

There are no guiding parts on the problems in this section, but we have provided hints that you can look up in the back of the book if you get stuck.

Problems

Problem 3.26: Determine $A + B$ if A and B are acute angles such that

$$\sin A + \sin B = \sqrt{\frac{3}{2}} \qquad \text{and} \qquad \cos A - \cos B = \sqrt{\frac{1}{2}}.$$

(Source: COMC) **Hints:** 84

Problem 3.27: Find the sum of the roots of $\tan^2 x - 9\tan x + 1 = 0$ that are between $x = 0$ and $x = 2\pi$ radians. *(Source: AHSME)* **Hints:** 137, 225

Problem 3.28: Let $a = \pi/2008$. Find the smallest positive integer n such that

$$2[\cos a \sin a + \cos 4a \sin 2a + \cos 9a \sin 3a + \cdots + \cos(n^2 a)\sin(na)]$$

is an integer. *(Source: AIME)* **Hints:** 189, 195

Problem 3.29: Let a, b, c, d be positive real numbers such that

$$a^2 + b^2 = 1,$$
$$c^2 + d^2 = 1,$$
$$ac - bd = \frac{1}{2}.$$

Calculate $ad + bc$. *(Source: Mandelbrot)* **Hints:** 79

Problem 3.30: Evaluate

$$\frac{\cos 1^\circ + \cos 2^\circ + \cos 3^\circ + \cdots + \cos 43^\circ + \cos 44^\circ}{\sin 1^\circ + \sin 2^\circ + \sin 3^\circ + \cdots + \sin 43^\circ + \sin 44^\circ}.$$

(Source: AIME) **Hints:** 103

Problem 3.31: Given any seven real numbers, prove that there are two of them, x and y, such that

$$0 \le \frac{x - y}{1 + xy} \le \frac{1}{\sqrt{3}}.$$

(Source: Canada) **Hints:** 212

Problem 3.32: Evaluate $(\sin 1^\circ)(\sin 3^\circ)(\sin 5^\circ)\cdots(\sin 177^\circ)(\sin 179^\circ)$. **Hints:** 29

Problem 3.26: Determine $A + B$ if A and B are acute angles such that

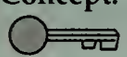
$$\sin A + \sin B = \sqrt{\frac{3}{2}} \qquad \text{and} \qquad \cos A - \cos B = \sqrt{\frac{1}{2}}.$$

(Source: COMC)

Solution for Problem 3.26: Our first instinct might be to try sum-to-product and difference-to-product identities since we have a sum of sines and a difference of cosines. Applying these identities to each equation gives

$$2\sin\left(\frac{A+B}{2}\right)\cos\left(\frac{A-B}{2}\right) = \sqrt{\frac{3}{2}},$$
$$-2\sin\left(\frac{A+B}{2}\right)\sin\left(\frac{A-B}{2}\right) = \sqrt{\frac{1}{2}}.$$

Uh-oh. It's not so clear what to do with that. We want $A + B$, but dividing the first equation by the second eliminates the terms involving $A + B$. It's not at all clear how we can get rid of the terms involving $A - B$. Let's look around for something different before trying to force this to work.

Concept:  Don't get locked into one way of attacking a problem—look for several starting points, go a little ways down each path, and then choose the one that looks most fruitful.

We expect that we'll have to combine the equations in some way, but it's not obvious that adding or multiplying them will do any good. We need some way to conveniently combine the sines and cosines, which brings us back to $\sin^2\theta + \cos^2\theta = 1$. The square roots on the right sides of the original equations are another reason we might think of this identity, since squaring each equation will get rid of the square roots. We're not sure what we'll do with the other terms that squaring produces, but maybe it will all work out in the end.

Squaring each equation gives:

$$\sin^2 A + 2\sin A\sin B + \sin^2 B = \frac{3}{2},$$
$$\cos^2 A - 2\cos A\cos B + \cos^2 B = \frac{1}{2}.$$

Adding these allows us to use $\sin^2 A + \cos^2 A = \sin^2 B + \cos^2 B = 1$, and we have

$$2 + 2\sin A\sin B - 2\cos A\cos B = 2,$$

so $\sin A\sin B - \cos A\cos B = 0$. This is promising, because we're looking for $A + B$, and the left side looks a lot like the identity for $\cos(A + B)$. Writing out that identity, we see that

$$\cos(A+B) = \cos A\cos B - \sin A\sin B = -(\sin A\sin B - \cos A\cos B),$$

and we just showed that the final expression on the right equals 0. Therefore, we have $\cos(A + B) = 0$. Since A and B are acute, we must have $0° < A + B < 180°$, so $A + B = 90°$. □

Taking a couple different first steps helped us avoid getting stuck. We weren't sure that either one of these would solve the problem, but they both initially looked promising. So, we took them both, and found that one of them gave us a result that we could use to solve the problem. (Our "failed" start will work, too, but it's quite a bit more work—see if you can figure out how.)

Concept: Experiment! Don't wait until you see the whole solution before you try something—you might end up waiting forever. Take little steps that might be promising, and reassess after each step.

This problem was also a good example of a very common strategy in dealing with trig identities:

Important: When stuck, try using $\sin^2\theta + \cos^2\theta = 1$.

Problem 3.27: Find the sum of the roots of $\tan^2 x - 9\tan x + 1 = 0$ that are between $x = 0$ and $x = 2\pi$ radians. *(Source: AHSME)*

Solution for Problem 3.27: Solution 1: Sines and cosines.

Concept: When in doubt in a trig problem involving functions other than sine and cosine, try converting everything to sine and cosine.

Writing the equation in terms of sine and cosine gives

$$\frac{\sin^2 x}{\cos^2 x} - \frac{9\sin x}{\cos x} + 1 = 0.$$

Multiplying both sides by $\cos^2 x$ gets rid of the fractions, and gives

$$\sin^2 x - 9\sin x\cos x + \cos^2 x = 0.$$

Aha! There's that $\sin^2 x + \cos^2 x$ again. Since this equals 1, we have $1 - 9\sin x\cos x = 0$, which means $\sin x\cos x = \frac{1}{9}$. We recognize the left side as $\frac{\sin 2x}{2}$, so our equation now is $\sin 2x = \frac{2}{9}$. What's wrong with this finish:

Bogus Solution: From $\sin 2x = \frac{2}{9}$, we have $2x = \arcsin\frac{2}{9}$, so the only solution to the original equation is $x = \frac{1}{2}\arcsin\frac{2}{9}$.

The problem here is that $\arcsin\frac{2}{9}$ is not the only value of $2x$ such that $\sin 2x = \frac{2}{9}$. What is wrong with this solution:

Bogus Solution: Angles whose sine is $\frac{2}{9}$ must have terminal points with y-coordinate $\frac{2}{9}$. There are two such points on the unit circle. One of these is the terminal point of $\arcsin\frac{2}{9}$ and the other is the terminal point of $\pi - \arcsin\frac{2}{9}$, since $\sin(\pi - \theta) = \sin\theta$. Therefore, in addition to $2x = \arcsin\frac{2}{9}$, which gives us $x = \frac{1}{2}\arcsin\frac{2}{9}$, we can also have $2x = \pi - \arcsin\frac{2}{9}$, which gives the solution $x = \frac{\pi}{2} - \frac{1}{2}\arcsin\frac{2}{9}$. The sum of these two solutions is $\frac{\pi}{2}$.

Alas, this is another convincing-but-wrong solution. It is true that there are only two angles from 0 to 2π whose sines equal $\frac{2}{9}$. However, we are told that the sine of $2x$ is $\frac{2}{9}$. Since $0 \le x \le 2\pi$, we have $0 \le 2x \le 4\pi$. Therefore, in addition to the two values of $2x$ we found in the second Bogus Solution, we also have two more between 2π and 4π. We get these values of $2x$ simply by adding 2π to the values we already found, since $\sin(\theta + 2\pi) = \sin\theta$ for all

θ. So, these extra two values of $2x$ are $2x = 2\pi + \arcsin\frac{2}{9}$ and $2x = 3\pi - \arcsin\frac{2}{9}$, which give $x = \pi + \frac{1}{2}\arcsin\frac{2}{9}$ and $x = \frac{3\pi}{2} - \frac{1}{2}\arcsin\frac{2}{9}$, respectively. The sum of these two solutions for x is $\frac{5\pi}{2}$, and adding this to the sum of the first two solutions we found gives a total of $\frac{5\pi}{2} + \frac{\pi}{2} = 3\pi$.

Solution 2: It's a quadratic in disguise! The given equation is quadratic in $\tan x$. To see this, we let $z = \tan x$, which gives us $z^2 - 9z + 1 = 0$. The quadratic formula then gives

$$\tan x = z = \frac{9 \pm \sqrt{81-4}}{2} = \frac{9 \pm \sqrt{77}}{2}.$$

Good luck finding the angles whose tangents equal those numbers without a calculator. We can at least tell that $\tan x$ is positive, which means that there are first and third quadrant angles that are solutions. But it doesn't appear to tell us much more. Maybe this is a dead end.

Looking back at the quadratic, we notice that the coefficients of the quadratic term and the constant term are both 1. This means that the product of the roots is 1. To see why, suppose the roots are r and s. Then, we can factor the quadratic as $(z-r)(z-s)$. Expanding this product gives $z^2 - (r+s)z + rs$. If this equals $z^2 - 9z + 1$, then we must have $rs = 1$.

But how does this help? All it tells us is that the roots of the equation are reciprocals. In other words, if $\tan\theta$ is one solution, then $1/\tan\theta$ is the other solution. But $1/\tan\theta$ is $\cot\theta$, so now we know that we can write the two solutions to $z^2 - 9z + 1 = 0$ as $\tan\theta$ and $\cot\theta$ for some value of θ. This doesn't tell us what θ is, but we don't need to know θ. We only need the sum of the possible values of θ.

Since $\cot\theta = \tan\left(\frac{\pi}{2} - \theta\right)$, we can now write the roots of $z^2 - 9z + 1 = 0$ as $\tan\theta$ and $\tan\left(\frac{\pi}{2} - \theta\right)$ for some value of θ. Success! We only want the sum of the values of x such that $\tan^2 x - 9\tan x + 1 = 0$. We just found that if $x = \theta$ is a solution, then so is $x = \frac{\pi}{2} - \theta$. Therefore, the two first quadrant solutions have sum $\frac{\pi}{2}$. Moreover, since $\tan(\pi + \theta) = \tan\theta$, we simply add π to θ and to $\frac{\pi}{2} - \theta$ to get the two third quadrant solutions, $\theta + \pi$ and $\frac{3\pi}{2} - \theta$. The sum of these is $\frac{5\pi}{2}$, so the sum of all four solutions is $\frac{\pi}{2} + \frac{5\pi}{2} = 3\pi$. □

Problem 3.28: Let $a = \pi/2008$. Find the smallest positive integer n such that

$$2[\cos a \sin a + \cos 4a \sin 2a + \cos 9a \sin 3a + \cdots + \cos(n^2 a)\sin(na)]$$

is an integer. *(Source: AIME)*

Solution for Problem 3.28: We don't see a good way to combine terms in the sum, so we turn to our product-to-sum identity for $\cos\alpha\sin\beta$, since that looks like the only tool we can use to get a different look at the problem. Applying this identity to first few terms gives

$$\cos a \sin a = \frac{1}{2}(\sin(a+a) + \sin(a-a)) = \frac{1}{2}(\sin 2a + \sin 0),$$
$$\cos 4a \sin 2a = \frac{1}{2}(\sin(2a+4a) + \sin(2a-4a)) = \frac{1}{2}(\sin 6a + \sin(-2a)),$$
$$\cos 9a \sin 3a = \frac{1}{2}(\sin(3a+9a) + \sin(3a-9a)) = \frac{1}{2}(\sin 12a + \sin(-6a)),$$
$$\cos 16a \sin 4a = \frac{1}{2}(\sin(4a+16a) + \sin(4a-16a)) = \frac{1}{2}(\sin 20a + \sin(-12a))$$

At first, adding these doesn't look promising. If only those negatives inside the sines in the last terms on the right were outside the sines instead—then we'd have a bunch of cancellation when we added.

Concept: Wishful thinking is a powerful problem solving tool—thinking about what you'd like to be true can focus your attention on the key step in a solution.

We don't have to just wish that $\sin(-a) = -\sin a$; this is an identity we already proved! So, now we can write the products in our sum as

$$\cos a \sin a = \frac{1}{2}(\sin 2a - \sin 0),$$
$$\cos 4a \sin 2a = \frac{1}{2}(\sin 6a - \sin 2a),$$
$$\cos 9a \sin 3a = \frac{1}{2}(\sin 12a - \sin 6a),$$
$$\cos 16a \sin 4a = \frac{1}{2}(\sin 20a - \sin 12a),$$

Adding all these will cancel out everything except $\frac{1}{2}\sin 20a$ and $-\frac{1}{2}\sin 0$. But will this cancellation continue? Will we always be able to cancel part of the k^{th} term with the term before it? We check by looking at what happens when we apply the product-to-sum identity to the $(k-1)^{\text{th}}$ and k^{th} terms:

$$\begin{aligned}\cos(k-1)^2 a \sin(k-1)a &= \frac{1}{2}\left(\sin(k-1+(k-1)^2)a + \sin(k-1-(k-1)^2)a\right)\\ &= \frac{1}{2}\left(\sin(k^2-k)a + \sin(-k^2+3k-2)a\right)\\ &= \frac{1}{2}(\sin k(k-1)a - \sin(k-1)(k-2)a),\\ \cos k^2 a \sin ka &= \frac{1}{2}\left(\sin(k+k^2)a + \sin(k-k^2)a\right) = \frac{1}{2}(\sin(k+1)ka - \sin k(k-1)a).\end{aligned}$$

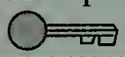
Sure enough, we have the expected cancellation when we add these terms, so we see that when we add the entire original series, the only terms that will not cancel are the $-\frac{1}{2}\sin 0$ from $\cos a \sin a$ and $\frac{1}{2}\sin(n+1)na$ from $\cos n^2 a \sin a$. Therefore, we have

$$2[\cos a \sin a + \cos 4a \sin 2a + \cos 9a \sin 3a + \cdots + \cos(n^2 a)\sin(na)] = 2\left(\frac{1}{2}\sin(n+1)na - \frac{1}{2}\sin 0\right) = \sin(n+1)na.$$

So, when $a = \pi/2008$, the original summation equals $\sin\frac{(n+1)n\pi}{2008}$. We seek the smallest positive value of n that makes this expression an integer. The sine of an angle is an integer if and only if the angle is an integer multiple of $\frac{\pi}{2}$. Therefore, $\frac{(n+1)n\pi}{2008} = \frac{k\pi}{2}$ for some integer k. Solving for k gives $k = \frac{(n+1)n}{1004}$, so 1004 must evenly divide $(n+1)n$. The prime factorization of 1004 is $2^2 \cdot 251$, so 251 must evenly divide $n+1$ or n. The smallest positive candidate for n is 250, but then 2^2 doesn't divide $(n+1)n$. The next smallest candidate, $n = 251$, works, since 2^2 divides 252. So, the smallest positive integer n that satisfies the problem is $n = 251$. □

A series in which part of each term cancels with part of another term in a way that allows us to write the series much more simply is called a **telescoping series**. A key step in our solution above is realizing that the sum could be written in a form that would "telescope" and leave us a much simpler expression.

Concept: Many series involving trigonometric functions can be written as telescoping series.



Problem 3.29: Let a, b, c, d be positive real numbers such that $a^2 + b^2 = 1$, $c^2 + d^2 = 1$, and $ac - bd = 1/2$. Calculate $ad + bc$. *(Source: Mandelbrot)*

Solution for Problem 3.29: The two equations $a^2 + b^2 = 1$ and $c^2 + d^2 = 1$ are in the same form as the identity $\sin^2\theta + \cos^2\theta = 1$. Since a, b, c, and d are positive real numbers that satisfy these two equations, we can make

the substitutions $a = \sin x$, $b = \cos x$, $c = \sin y$, and $d = \cos y$, where x and y are first quadrant angles. With this substitution, we are guaranteed to have $a^2 + b^2 = c^2 + d^2 = 1$. Let's see what this substitution gives us in the third equation:

$$(\sin x)(\sin y) - (\cos x)(\cos y) = \frac{1}{2}.$$

We now recognize a trig identity in the third equation as well; we can write the left side as

$$(\sin x)(\sin y) - (\cos x)(\cos y) = -((\cos x)(\cos y) - (\sin x)(\sin y)) = -\cos(x + y),$$

so the third equation now is $-\cos(x + y) = \frac{1}{2}$, which gives us $\cos(x + y) = -\frac{1}{2}$. Since x and y are first quadrant angles, we know that $x + y$ is in the first or second quadrant. The only such angle for which $\cos(x + y) = -\frac{1}{2}$ is $120°$.

But how does this help with $ad + bc$? Let's see what our substitution does to this expression:

$$ad + bc = (\sin x)(\cos y) + (\cos x)(\sin y).$$

Another trig identity! We have

$$(\sin x)(\cos y) + (\cos x)(\sin y) = \sin(x + y).$$

Since $x + y = 120°$, we have $ad + bc = \sin 120° = \frac{\sqrt{3}}{2}$. □

Concept: Trigonometry can be a powerful tool even in problems that don't appear to be about trigonometry.

A key sign that a trigonometric substitution might be helpful is the presence of an expression or equation that is conspicuously in a form found in a common trig identity.

Problem 3.30: Evaluate

$$\frac{\cos 1° + \cos 2° + \cos 3° + \cdots + \cos 43° + \cos 44°}{\sin 1° + \sin 2° + \sin 3° + \cdots + \sin 43° + \sin 44°}.$$

(Source: AIME)

Solution for Problem 3.30: The sum of trig functions suggests the sum-to-product identities, but there are 44 cosines in the numerator—to which pairs should we apply the identities?

Concept: Look for convenient relationships among the angles in trig problems.

The angles can be paired off into sums that equal $45°$, which is an angle we know how to handle. Applying the cosine sum-to-product identity to these pairs gives

$$\cos 44° + \cos 1° = 2\cos\left(\frac{44° + 1°}{2}\right)\cos\left(\frac{44° - 1°}{2}\right) = 2\cos\frac{45°}{2}\cos\frac{43°}{2},$$

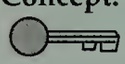
$$\cos 43° + \cos 2° = 2\cos\left(\frac{43° + 2°}{2}\right)\cos\left(\frac{43° - 2°}{2}\right) = 2\cos\frac{45°}{2}\cos\frac{41°}{2},$$

$$\cos 42° + \cos 3° = 2\cos\left(\frac{42° + 3°}{2}\right)\cos\left(\frac{42° - 3°}{2}\right) = 2\cos\frac{45°}{2}\cos\frac{39°}{2},$$

$$\vdots$$

$$\cos 23° + \cos 22° = 2\cos\left(\frac{23° + 22°}{2}\right)\cos\left(\frac{23° - 22°}{2}\right) = 2\cos\frac{45°}{2}\cos\frac{1°}{2}.$$

And when we add all these up, we get... quite a mess. We can handle $\cos\frac{45}{2}^\circ$ with the half angle identity, but the rest of those cosines are intimidating.

Concept:  When you encounter a nasty complication in a trig problem, and can't find an identity to wipe it out, look for something else in the problem to do the dirty work for you.

We could try pairing the cosines off again, but that will generate even more angles we don't know how to handle. (Try it and see.) Let's at least take a look at what happens with the sines in the denominator when we apply the sum-to-product identities. Maybe it will all work out in the end:

$$\begin{aligned}
\sin 44^\circ + \sin 1^\circ &= 2\sin\left(\frac{44^\circ+1^\circ}{2}\right)\cos\left(\frac{44^\circ-1^\circ}{2}\right) = 2\sin\frac{45^\circ}{2}\cos\frac{43^\circ}{2},\\
\sin 43^\circ + \sin 2^\circ &= 2\sin\left(\frac{43^\circ+2^\circ}{2}\right)\cos\left(\frac{43^\circ-2^\circ}{2}\right) = 2\sin\frac{45^\circ}{2}\cos\frac{41^\circ}{2},\\
\sin 42^\circ + \sin 3^\circ &= 2\sin\left(\frac{42^\circ+3^\circ}{2}\right)\cos\left(\frac{42^\circ-3^\circ}{2}\right) = 2\sin\frac{45^\circ}{2}\cos\frac{39^\circ}{2},\\
&\vdots\\
\sin 23^\circ + \sin 22^\circ &= 2\sin\left(\frac{23^\circ+22^\circ}{2}\right)\cos\left(\frac{23^\circ-22^\circ}{2}\right) = 2\sin\frac{45^\circ}{2}\cos\frac{1^\circ}{2}.
\end{aligned}$$

There are those cosines again, but this time we're happy to see them! We can use these to cancel out the ones we found when we paired off the cosines. We have

$$\frac{\cos 1^\circ + \cos 2^\circ + \cdots + \cos 44^\circ}{\sin 1^\circ + \sin 2^\circ + \cdots + \sin 44^\circ} = \frac{2\cos\frac{45}{2}^\circ\left(\cos\frac{43}{2}^\circ + \cos\frac{41}{2}^\circ + \cdots + \cos\frac{1}{2}^\circ\right)}{2\sin\frac{45}{2}^\circ\left(\cos\frac{43}{2}^\circ + \cos\frac{41}{2}^\circ + \cdots + \cos\frac{1}{2}^\circ\right)} = \cot\frac{45^\circ}{2}.$$

Now, we know what to do. We apply the half angle identity for tangent to find

$$\cot\frac{45^\circ}{2} = \frac{1}{\tan\frac{45^\circ}{2}} = \frac{1}{(\sin 45^\circ)/(1+\cos 45^\circ)} = \frac{1+\cos 45^\circ}{\sin 45^\circ} = \frac{1+\frac{\sqrt{2}}{2}}{\frac{\sqrt{2}}{2}} = \frac{2+\sqrt{2}}{\sqrt{2}} = \sqrt{2}+1.$$

□

Problem 3.31: Given any seven real numbers, prove that there are two of them, x and y, such that

$$0 \le \frac{x-y}{1+xy} \le \frac{1}{\sqrt{3}}.$$

(Source: Canada)

Solution for Problem 3.31: There don't appear to be any good algebraic tools for attacking this problem. Even multiplying both sides of the right side inequality by $1+xy$ is complicated by the fact that $1+xy$ might be negative. However, the expression $(x-y)/(1+xy)$ does resemble the angle difference identity for tangent:

$$\tan(A-B) = \frac{\tan A - \tan B}{1+\tan A\tan B}.$$

Another clue that we can use this identity is that x and y can be "any real numbers." The range of $\tan\theta$ is all real numbers, so letting $x = \tan A$ and $y = \tan B$ is a permissible substitution, since no matter what x and y are, there

are angles A and B such that $x = \tan A$ and $y = \tan B$. Writing the given inequality with this substitution, we have

$$0 \le \frac{\tan A - \tan B}{1 + \tan A \tan B} \le \frac{1}{\sqrt{3}}.$$

Applying the angle difference identity for tangent, we have

$$0 \le \tan(A - B) \le \frac{1}{\sqrt{3}}.$$

What does this tell us about A and B? The algebraic expression is not the only part of the problem we can relate to trig. Viewing the two constants on the ends of the inequality as tangents of angles as well, we have

$$\tan 0 \le \tan(A - B) \le \tan \frac{\pi}{6}.$$

Now we have a different view of the problem. We can view each of the seven real numbers as the tangent of some angle from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$, and our problem becomes:

> Given any seven angles whose measures are in the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, prove that there are two of them, A and B, such that
>
> $$\tan 0 \le \tan(A - B) \le \tan \frac{\pi}{6}.$$

We now consider what angles have tangents between $\tan 0$ and $\tan \frac{\pi}{6}$. We have $\tan 0 \le \tan(A-B) \le \tan \frac{\pi}{6}$ when the slope of the line through the origin and the terminal point of $A - B$ is positive but no greater than the slope of the line through the origin and the terminal point of $\frac{\pi}{6}$. The only first quadrant angles for which this is true are $0 \le A - B \le \frac{\pi}{6}$. (It is also true for some third quadrant angles, but because A and B are in the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, the angle $A - B$ cannot be in the third quadrant.)

All we have to do now is show that for any seven angles in the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, there are two that are no more than $\frac{\pi}{6}$ apart. To do this, we note that we can break the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ into 6 non-overlapping intervals that are no longer than $\frac{\pi}{6}$:

$$\left(-\frac{\pi}{2}, -\frac{\pi}{3}\right]; \quad \left(-\frac{\pi}{3}, -\frac{\pi}{6}\right]; \quad \left(-\frac{\pi}{6}, 0\right]; \quad \left(0, \frac{\pi}{6}\right]; \quad \left(\frac{\pi}{6}, \frac{\pi}{3}\right]; \quad \left(\frac{\pi}{3}, \frac{\pi}{2}\right).$$

If any two angles are in the same interval, then they are no more than $\frac{\pi}{6}$ apart, and we have the two desired angles. If we restrict ourselves to one angle per interval, the most we can have is 6 angles. Since we have 7 angles distributed among these 6 intervals, there must be an interval with at least 2 angles. Letting A be the larger of these two angles and B be the smaller, we have $0 \le A - B \le \frac{\pi}{6}$, so $\tan 0 \le \tan(A - B) \le \tan \frac{\pi}{6}$, as desired. □

Sidenote: A key step in our solution to Problem 3.31 is the observation that if 7 angles are distributed among 6 intervals, then at least 2 of the angles must be in the same interval. This is an example of the **Pigeonhole Principle**, which states that if more than n objects are distributed among n holes, then at least one hole must contain more than one of the objects. We discuss the Pigeonhole Principle in more detail in Art of Problem Solving's *Intermediate Counting & Probability*.

Problem 3.32: Evaluate $(\sin 1°)(\sin 3°)(\sin 5°)\cdots(\sin 177°)(\sin 179°)$.

Solution for Problem 3.32: We can pair off terms whose angles sum to 180°. We might try the product-to-sum identity (give it a try), but seeing the 180° suggests that we can also use $\sin(180° - \theta) = \sin\theta$ to give

$$(\sin 1°)(\sin 3°)(\sin 5°)\cdots(\sin 177°)(\sin 179°) = (\sin 1°)(\sin 3°)\cdots(\sin 89°)(\sin 89°)(\sin 87°)\cdots(\sin 3°)(\sin 1°).$$

We can now pair up the sines as squares of sines. This suggests trying $\sin^2\theta = 1 - \cos^2\theta$, but that turns our product into a product of differences, which looks scary. If only those products of two sines of the same angle were the product of the sine and the cosine of the same angle, then we could use the sine double angle identity in reverse to simplify the problem.

This bit of wishful thinking pays off when we recognize that we can again pair off the angles in a convenient way, since we can pair them such that the angles sum to 90°. We want to turn some sines into cosines and we have angles that add to 90°, so the identity $\sin\theta = \cos(90° - \theta)$ might help us out. Applying this to the second batch of sines in the product gives:

$$\begin{aligned}(\sin 1°)(\sin 3°)\cdots(\sin 89°)(\sin 89°)(\sin 87°)\cdots(\sin 3°)(\sin 1°)&\\ = (\sin 1°)(\sin 3°)\cdots(\sin 89°)(\cos 1°)(\cos 3°)\cdots(\cos 87°)(\cos 89°).&\end{aligned}$$

Notice that we could have gotten straight to this from this initial product by applying $\sin\theta = \cos(\theta - 90°)$.

Concept: There are many ways to tackle most trig identity problems. You won't always find the fastest way, and reviewing your solutions may reveal relationships that will help you solve future problems.

We can now apply the sine double angle identity in reverse:

$$\begin{aligned}(\sin 1°)(\sin 3°)\cdots(\sin 89°)(\cos 1°)(\cos 3°)\cdots(\cos 89°) &= (\sin 1°)(\cos 1°)(\sin 3°)(\cos 3°)\cdots(\sin 89°)(\cos 89°)\\ &= \left(\frac{\sin 2°}{2}\right)\left(\frac{\sin 6°}{2}\right)\cdots\left(\frac{\sin 178°}{2}\right)\\ &= \frac{1}{2^{45}}(\sin 2°)(\sin 6°)\cdots(\sin 178°).\end{aligned}$$

Once again, we have angles that pair to 180°, so we try $\sin\theta = \sin(180° - \theta)$. But we have to be careful:

WARNING!! Watch out for special cases. When pairing off terms in a sequence, a summation, or a product, you have to pay special attention to the middle term if there are an odd number of terms.

Here, we have 45 terms, so we can't pair them all off. There's one extra term—the middle term, which is $\sin 90°$. But this simply equals 1, so we can continue without it, since it doesn't affect the product. After applying the identity $\sin(180° - \theta) = \sin\theta$ to all terms with angles greater than 90°, we have

$$\frac{1}{2^{45}}(\sin 2°)(\sin 6°)\cdots(\sin 178°) = \frac{1}{2^{45}}(\sin 2°)(\sin 2°)(\sin 6°)(\sin 6°)\cdots(\sin 86°)(\sin 86°).$$

Uh-oh. We can't pair these angles off to add to 90° like we did before. If we apply $\sin\theta = \cos(90° - \theta)$ to the second sine in each pair, we have

$$\frac{1}{2^{45}}(\sin 2°)(\cos 88°)(\sin 6°)(\cos 84°)\cdots(\sin 86°)(\cos 4°).$$

Now what? Looks like we're stuck.

Concept: Don't get so mired in one-step tactics that you stop thinking strategically. When hopelessly stuck, step back and think more broadly about what you're trying to accomplish.

We'd like to make the huge product

$$\frac{1}{2^{45}}(\sin 2°)(\sin 2°)(\sin 6°)(\sin 6°)\cdots(\sin 86°)(\sin 86°)$$

collapse in some way. We collapsed a sum in Problem 3.28 by writing each term as a difference, so that part of each term cancels with another. To do the same in a product, we write each term as a quotient, hoping the denominator of each term will cancel with the numerator of another. So, we look for ways to rewrite some of our terms as quotients. A natural candidate for this approach is the sine double angle formula, because we've used it fruitfully this way before.

We write $\sin 2° = \frac{\sin 4°}{2\cos 2°}$, and think about whether or not doing the same for other terms will produce the desired cancellation. It's not immediately obvious that the $\cos 2°$ will cancel with a numerator in another term, but the $\sin 4°$ makes us focus on the $\sin 86°$ term, since $\cos 86° = \sin 4°$. Indeed, we have $\sin 86° = \frac{\sin 172°}{2\cos 86°}$, and we see that we'll get some cancellation. So, rewriting all the terms in our product similarly, we have:

$$\frac{1}{2^{45}}(\sin 2°)(\sin 2°)(\sin 6°)(\sin 6°)\cdots(\sin 86°)(\sin 86°) = \frac{1}{2^{45}}\left(\frac{\sin 4°}{2\cos 2°}\right)^2\left(\frac{\sin 12°}{2\cos 6°}\right)^2\cdots\left(\frac{\sin 172°}{2\cos 86°}\right)^2.$$

We then apply $\cos\theta = \sin(90° - \theta)$, knowing we will get some cancellation. We find:

$$\frac{1}{2^{45}}\left(\frac{\sin 4°}{2\cos 2°}\right)^2\left(\frac{\sin 12°}{2\cos 6°}\right)^2\cdots\left(\frac{\sin 172°}{2\cos 86°}\right)^2 = \frac{1}{2^{45}}\cdot\frac{1}{2^{44}}\left(\frac{\sin 4°}{\sin 88°}\right)^2\left(\frac{\sin 12°}{\sin 84°}\right)^2\left(\frac{\sin 20°}{\sin 80°}\right)^2\cdots\left(\frac{\sin 172°}{\sin 4°}\right)^2.$$

We see that some of the sines immediately cancel, but some are left over. The sine of each multiple of 4° less than 90° appears in a denominator. The numerator contains sines of every multiple of 4° less than 180° that is not a multiple of 8°. Cancelling the common sines leaves only sines of multiples of 8° in the denominator and the sines of angles greater than 90° in the numerator:

$$\frac{1}{2^{45}}\cdot\frac{1}{2^{44}}\left(\frac{\sin 4°}{\sin 88°}\right)^2\left(\frac{\sin 12°}{\sin 84°}\right)^2\left(\frac{\sin 20°}{\sin 80°}\right)^2\cdots\left(\frac{\sin 172°}{\sin 4°}\right)^2 = \frac{1}{2^{89}}\left(\frac{(\sin 92°)(\sin 100°)(\sin 108°)\cdots(\sin 172°)}{(\sin 88°)(\sin 80°)(\sin 72°)\cdots(\sin 8°)}\right)^2.$$

And now, the identity $\sin\theta = \sin(180° - \theta)$ provides the finishing touch:

$$\frac{1}{2^{89}}\left(\frac{(\sin 92°)(\sin 100°)(\sin 108°)\cdots(\sin 172°)}{(\sin 88°)(\sin 80°)(\sin 72°)\cdots(\sin 8°)}\right)^2 = \frac{1}{2^{89}}\left(\frac{(\sin 92°)(\sin 100°)(\sin 108°)\cdots(\sin 172°)}{(\sin 92°)(\sin 100°)(\sin 108°)\cdots(\sin 172°)}\right)^2 = \frac{1}{2^{89}}.$$

The final result suggests that there might be a better way to solve this problem—there are 90 terms in the original product and the answer is $2\cdot\frac{1}{2^{90}}$. Let's look for a better solution.

Concept: If you do a lot of arduous work on a problem and come to a relatively simple answer, try to use that answer as a guide to find a simpler solution.

So, we go back to the original product:

$$(\sin 1°)(\sin 3°)(\sin 5°)\cdots(\sin 177°)(\sin 179°).$$

We'd like to find a way to introduce 89 or 90 factors of $\frac{1}{2}$. We've seen how to introduce factors of $\frac{1}{2}$ with the double angle formula for sine, such as when we wrote $\sin 2° = \frac{\cos 2°}{2\sin 4°}$. We expect that this means we'll have to introduce

terms with an even number of degrees, since we'll be using the sine double angle formula. A natural choice now is to introduce all the missing even degree measures by writing the desired expression as a quotient:

$$(\sin 1^\circ)(\sin 3^\circ)(\sin 5^\circ)\cdots(\sin 177^\circ)(\sin 179^\circ) = \frac{(\sin 1^\circ)(\sin 2^\circ)(\sin 3^\circ)\cdots(\sin 178^\circ)(\sin 179^\circ)}{(\sin 2^\circ)(\sin 4^\circ)(\sin 6^\circ)\cdots(\sin 176^\circ)(\sin 178^\circ)}.$$

This gives us 89 terms in the denominator, which suggests we're on the right track. We need to pair up factors in the numerator in a way that allows us to finish with the sine double angle formula. Therefore, we need to find $\cos 1^\circ, \cos 2^\circ, \ldots, \cos 89^\circ$ to pair up with $\sin 1^\circ, \sin 2^\circ, \ldots, \sin 89^\circ$. This means we have to focus on turning the sines of obtuse angles into cosines of acute angles. The identity $\sin\theta = \cos(\theta - 90^\circ)$ does the trick, and we have

$$\begin{aligned}
\frac{(\sin 1^\circ)(\sin 2^\circ)(\sin 3^\circ)\cdots(\sin 178^\circ)(\sin 179^\circ)}{(\sin 2^\circ)(\sin 4^\circ)(\sin 6^\circ)\cdots(\sin 176^\circ)(\sin 178^\circ)} &= \frac{(\sin 1^\circ)(\sin 2^\circ)\cdots(\sin 89^\circ)(\sin 90^\circ)(\cos 1^\circ)(\cos 2^\circ)\cdots(\cos 89^\circ)}{(\sin 2^\circ)(\sin 4^\circ)(\sin 6^\circ)\cdots(\sin 176^\circ)(\sin 178^\circ)} \\
&= \frac{(\sin 1^\circ)(\cos 1^\circ)(\sin 2^\circ)(\cos 2^\circ)\cdots(\sin 89^\circ)(\cos 89^\circ)(1)}{(\sin 2^\circ)(\sin 4^\circ)(\sin 6^\circ)\cdots(\sin 176^\circ)(\sin 178^\circ)} \\
&= \frac{\left(\frac{1}{2}\sin 2^\circ\right)\left(\frac{1}{2}\sin 42^\circ\right)\cdots\left(\frac{1}{2}\sin 178^\circ\right)}{(\sin 2^\circ)(\sin 4^\circ)(\sin 6^\circ)\cdots(\sin 176^\circ)(\sin 178^\circ)} \\
&= \frac{1}{2^{89}}.
\end{aligned}$$

That's much easier, but it would have been a pretty hard solution to find without the big clue that we're looking for 89 factors of $\frac{1}{2}$. With that clue in mind, we found a new strategy:

Concept: Some products can be evaluated by strategically multiplying by expressions that equal 1 in order to introduce convenient cancellation.

We did so in our second solution by multiplying by a bunch of factors of the form $\frac{\sin 2n^\circ}{\sin 2n^\circ}$. □

Our solution to Problem 3.32 is a good example of how to tackle hard problems. We didn't have a single magical observation that solved the problem. We often took steps without knowing for sure that they'd lead to a solution. We used wishful thinking to help guide us, and stepped back when one of our forward steps proved to be a blind alley. Above all, we didn't give up—we kept trying new approaches until we found a combination of tactics that solved the problem. And after we solved the problem, we used the answer, together with some other clues in our solution, to go back and find a nicer solution.

Exercises

3.5.1 Find A such that $0^\circ < A < 90^\circ$ and $\cos 41^\circ + \sin 41^\circ = \sqrt{2}\sin A$. *(Source: NYSML)*

3.5.2 Find a solution for Problem 3.29 that doesn't require trigonometry.

3.5.3 Determine the sum of the angles A and B, where $0^\circ \le A, B \le 180^\circ$, and

$$\sin A + \sin B = \sqrt{\frac{3}{2}}, \quad \cos A + \cos B = \sqrt{\frac{1}{2}}.$$

(Source: COMC) (Note: This is *not* exactly the same as Problem 3.26.)

3.5.4 If $\tan x + \tan y = 25$ and $\cot x + \cot y = 30$, then what is $\tan(x + y)$? *(Source: AIME)*

3.5.5 Let A and B be acute angles such that $\tan A = 1/7$ and $\sin B = 1/\sqrt{10}$. Find the degree measure of $A + 2B$ without using a calculator.

3.5.6 Find all acute angles θ such that $\sin\theta + \sin 2\theta = \cos\theta + \cos 2\theta$.

3.5.7 Given $\cos(\alpha+\beta) + \sin(\alpha-\beta) = 0$ and $\tan\beta = 1/2000$, find $\tan\alpha$. *(Source: HMMT)*

3.5.8★ Given that

$$(1+\sin t)(1+\cos t) = \frac{5}{4},$$

find $(1-\sin t)(1-\cos t)$. *(Source: AIME)* **Hints:** 67, 44

3.5.9★ Solve the system

$$\begin{aligned} 2x + yx^2 &= y, \\ 2y + zy^2 &= z, \\ 2z + xz^2 &= x. \end{aligned}$$

Hints: 240

3.6 Summary

In this chapter, we explored the following identities:

$$\begin{array}{llll} \sin^2\theta + \cos^2\theta = 1 & \sin(-\theta) = -\sin\theta & \sin(90^\circ-\theta) = \cos\theta & \sin(180^\circ-\theta) = \sin\theta \\ \tan^2\theta + 1 = \sec^2\theta & \cos(-\theta) = \cos\theta & \cos(90^\circ-\theta) = \sin\theta & \cos(180^\circ-\theta) = -\cos\theta \\ \cot^2\theta + 1 = \csc^2\theta & \tan(-\theta) = -\tan\theta & \tan(90^\circ-\theta) = \cot\theta & \tan(180^\circ-\theta) = -\tan\theta \end{array}$$

Angle sum and difference identities:

$$\begin{aligned} \sin(\alpha+\beta) &= \sin\alpha\cos\beta + \sin\beta\cos\alpha & \sin(\alpha-\beta) &= \sin\alpha\cos\beta - \sin\beta\cos\alpha \\ \cos(\alpha+\beta) &= \cos\alpha\cos\beta - \sin\alpha\sin\beta & \cos(\alpha-\beta) &= \cos\alpha\cos\beta + \sin\alpha\sin\beta \\ \tan(\alpha+\beta) &= \frac{\tan\alpha+\tan\beta}{1-\tan\alpha\tan\beta} & \tan(\alpha-\beta) &= \frac{\tan\alpha-\tan\beta}{1+\tan\alpha\tan\beta} \end{aligned}$$

Double angle and half-angle identities:

$$\begin{aligned} \sin 2x &= 2\sin x\cos x & \cos\frac{\theta}{2} &= \pm\sqrt{\frac{1+\cos\theta}{2}} \\ \cos 2x &= \cos^2 x - \sin^2 x = 2\cos^2 x - 1 = 1 - 2\sin^2 x & \sin\frac{\theta}{2} &= \pm\sqrt{\frac{1-\cos\theta}{2}} \\ \tan 2x &= \frac{2\tan x}{1-\tan^2 x} & \tan\frac{\theta}{2} &= \frac{\sin\theta}{1+\cos\theta} = \frac{1-\cos\theta}{\sin\theta} \end{aligned}$$

Product-to-sum and sum-to-product identities:

$$\begin{aligned} \cos\alpha\cos\beta &= \frac{1}{2}\left(\cos(\alpha+\beta) + \cos(\alpha-\beta)\right) & \cos x + \cos y &= 2\cos\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right) \\ \sin\alpha\sin\beta &= \frac{1}{2}\left(\cos(\alpha-\beta) - \cos(\alpha+\beta)\right) & \cos x - \cos y &= -2\sin\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right) \\ \sin\alpha\cos\beta &= \frac{1}{2}\left(\sin(\alpha+\beta) + \sin(\alpha-\beta)\right) & \sin x + \sin y &= 2\sin\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right) \\ & & \sin x - \sin y &= 2\sin\left(\frac{x-y}{2}\right)\cos\left(\frac{x+y}{2}\right) \end{aligned}$$

Things To Watch Out For!

WARNING!! Whenever you solve an equation by performing some possibly irreversible steps, you must check your solutions at the end, to make sure they satisfy the original equation. Two common irreversible steps are squaring an equation and multiplying both sides of an equation by an expression that might equal zero.

Problem Solving Strategies

Concepts:

- If a tactic makes some progress on a problem, but doesn't entirely solve it, look for a way to use the tactic again.
- Don't get locked into one way of attacking a problem—look for several starting points, go a little ways down each path, and then choose the one that looks most fruitful.
- Experiment! Don't wait until you see the whole solution before you try something—you might end up waiting forever. Take little steps that might be promising, and reassess after each step.
- Wishful thinking is a powerful problem solving tool—thinking about what you'd like to be true can focus your attention on the key step in a solution.
- A fruitful first step in many trig problems is rewriting the problem in terms of sine and cosine, because most people have more experience with sine and cosine than with the other trig functions.
- When solving an equation that contains trig functions, it's often best to write the equation in terms of a single trig function, so that you can then try to solve the equation for that function.
- Recognizing the forms of common trigonometric identities can help solve many trigonometric problems.
- Recognizing relationships among angles in a complicated trig expression can help simplify the expression.
- Many series involving trigonometric functions can be written as telescoping series.
- Trigonometry can be a powerful tool even in problems that don't appear to be about trigonometry.

Extra! You may have heard that prior to calculators, people sometimes used logarithm tables to perform multiplication and division of large numbers. But what did they do before mathematicians developed logarithms? They used trig! More specifically, they used the product-to-sum and product-to-difference identities to turn products into easier-to-handle sums and differences. If you can't figure out how this works, look up **prosthaphaeresis** on the internet.

REVIEW PROBLEMS

3.33 Prove that $\tan(\pi - \theta) = -\tan\theta$.

3.34 Express $\sin(\theta - \pi)$ in terms of $\cos\theta$ and/or $\sin\theta$.

3.35 Prove each of the identities below using only the identities $\sin(-x) = \sin x$, $\cos(-x) = -\cos x$, $\sin(90° - x) = \cos x$, $\sin^2 x + \cos^2 x = 1$, and $\sin(x + y) = \sin x \cos y + \sin y \cos x$. Try to do this problem without looking back in the book. The goal of this problem is to learn that you do not have to memorize all of these identities—you can rebuild them from a small group of basic identities whenever you need them.

(a) $\cos(x - y) = \cos x \cos y + \sin x \sin y$

(b) $\sin 2x = 2\sin x \cos x$

(c) $\tan(x + y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}$

(d) $\cos\frac{x}{2} = \sqrt{\frac{1+\cos x}{2}}$ for $0 < x < \pi$

(e) $\cos x \cos y = \frac{1}{2}(\cos(x + y) + \cos(x - y))$

(f) $\sin x - \sin y = 2\sin\left(\frac{x-y}{2}\right)\cos\left(\frac{x+y}{2}\right)$

3.36 Simplify $\sin(x - y)\cos y + \cos(x - y)\sin y$. *(Source: AHSME)*

3.37 Which of the following equals $\sqrt{2}\cos\left(\frac{\pi}{4} - x\right)$?

(a) $\cos x + \sin x$

(b) $-\cos x + \sin x$

(c) $\cos x - \sin x$

(d) $-\cos x - \sin x$

3.38 Determine the constants a and b so that $\dfrac{-3 + 4\cos^2\theta}{1 - 2\sin\theta} = a + b\sin\theta$. *(Source: CEMC)*

3.39 What is the range of the function $g(t) = 2\sin t + 3\cos t$?

3.40 Compute $\dfrac{\tan^2 20° - \sin^2 20°}{\tan^2 20° \sin^2 20°}$. *(Source: HMMT)*

3.41 Show that $\dfrac{\tan(\alpha - \beta) + \tan\beta}{1 - \tan(\alpha - \beta)\tan\beta} = \tan\alpha$.

3.42 Express $\tan\left(\frac{3\pi}{2} + \theta\right)$ in terms of $\tan\theta$.

3.43 Show that there is no triangle ABC for which $\cos A \cos B - \sin A \sin B = 1$.

3.44 The complicated trigonometric expression shown below is equivalent to either $\sin x$, $\cos x$, or $\tan x$. Which one is it?

$$\frac{\sqrt{3}\sin(x + 30°) - \cos(x + 30°)}{4\cos x \sin(x + 30°) - 4\sin x \cos(x + 30°)}.$$

(Source: Mandelbrot)

3.45 Evaluate $\tan\left(\arccos\left(-\frac{1}{2}\right) + \arcsin\frac{\sqrt{2}}{2}\right)$.

3.46 Compute $(\sin 15° + \sin 75°)^6$. *(Source: ARML)*

3.47 Simplify $\dfrac{\sin 2x}{\sin x} - \dfrac{\cos 2x}{\cos x}$.

3.48 Find all possible values of $\sin 2\theta$ if $\cos\theta = -0.6$.

3.49 Show that $\dfrac{\cot x - 1}{\cot x + 1} = \dfrac{\cos 2x}{1 + \sin 2x}$.

3.50 If $\sin 2x = 24/25$, find the value of $\sin^4 x + \cos^4 x$.

3.51 Find $\tan \frac{5\pi}{8}$.

3.52 Find all x such that $0 \le x \le 2\pi$ and $\sin 2x \le \sin x$.

3.53 For $-\pi \le \theta \le \pi$, find all solutions to the equation $2(\sin^2 \theta - \cos^2 \theta) = 8 \sin \theta - 5$. *(Source: CEMC)*

3.54

(a) Show that $\dfrac{\sin \theta + \sin 2\theta}{1 + \cos \theta + \cos 2\theta} = \tan \theta$ if $0 \le \theta < \pi/2$.

(b) If $\pi/2 < \theta < \pi$, for what value(s) of θ is the identity in part (a) not true? *(Source: CEMC)*

3.55 Find the acute angle ϕ such that $2 \sin 52^\circ \sin 68^\circ = \frac{1}{2} + \sin \phi$.

3.56 Prove the identity $\dfrac{\tan x}{1 - \cot x} + \dfrac{\cot x}{1 - \tan x} = \sec x \csc x + 1$.

3.57 Show that $\dfrac{\sin x + \sin 3x + \sin 5x}{\cos x + \cos 3x + \cos 5x} = \tan 3x$ for all x for which $\tan 3x$ is defined.

Challenge Problems

3.58 Find the value of $\tan x$ if $\sin x + \cos x = 1/5$ and $\pi/2 < x < \pi$. **Hints:** 44

3.59 If $f\left(\frac{x}{x-1}\right) = \frac{1}{x}$ for all $x \neq 0, 1$, and $0 < \theta < \pi/2$, then find $f(\sec^2 \theta)$. **Hints:** 156

3.60 Compute the smallest positive angle x, in degrees, such that $\tan 4x = \dfrac{\cos x - \sin x}{\cos x + \sin x}$. *(Source: NYSML)* **Hints:** 261

3.61 Determine all points at which the graphs of $y = 8 \cos x + 5 \tan x$ and $y = 5 \sec x + 2 \sin 2x$ intersect, where $0 \le x \le 2\pi$. *(Source: CEMC)*

3.62 Find all x between $-\pi/2$ and $\pi/2$ such that $1 - \sin^4 x - \cos^2 x = 1/16$. *(Source: HMMT)*

3.63 Let y be a fixed real number, $0 < y < \pi$, and let $f(x) = \dfrac{\sin x + \sin(x + y)}{\cos x - \cos(x + y)}$ for $0 \le x \le \pi - y$. Show that f is constant. **Hints:** 166

3.64 Suppose that $\sec x + \tan x = 22/7$. Find $\csc x + \cot x$. *(Source: AIME)* **Hints:** 147

3.65 Prove that for all real x and y, we have $-\dfrac{1}{2} \le \dfrac{(x + y)(1 - xy)}{(1 + x^2)(1 + y^2)} \le \dfrac{1}{2}$. **Hints:** 253

3.66 Determine all θ such that $0 \le \theta \le \frac{\pi}{2}$ and $\sin^5 \theta + \cos^5 \theta = 1$. **Hints:** 122

3.67 Find, with proof, all real numbers x between 0 and 2π such that $\tan 7x - \sin 6x = \cos 4x - \cot 7x$. *(Source: USAMTS)* **Hints:** 267

3.68 Find the value of $10 \cot(\cot^{-1} 3 + \cot^{-1} 7 + \cot^{-1} 13 + \cot^{-1} 21)$. *(Source: AIME)* **Hints:** 26

3.69 Find $\cos \frac{\pi}{7} \cdot \cos \frac{2\pi}{7} \cdot \cos \frac{4\pi}{7}$. **Hints:** 223

3.70

(a) Express $\tan 3x$ in terms of $\tan x$.

(b) Let θ be the angle in the second quadrant for which $\cos\theta = -8/17$. Determine $\tan(3\theta/2)$. *(Source: Mandelbrot)*

(c) Solve for x: $\tan 20^\circ \tan 40^\circ \tan 80^\circ = \tan x$. **Hints:** 72, 20

3.71 If $\sin\theta + \cos\theta + \tan\theta + \cot\theta + \sec\theta + \csc\theta = 7$, then find $\sin 2\theta$. *(Source: NYSML)* **Hints:** 99, 168

3.72★ The first two terms of a sequence are $a_1 = 1$ and $a_2 = \frac{1}{\sqrt{3}}$. For $n \geq 1$,

$$a_{n+2} = \frac{a_n + a_{n+1}}{1 - a_n a_{n+1}}.$$

What is $|a_{2009}|$? *(Source: AMC)* **Hints:** 208

3.73★ Find the least positive integer n such that

$$\frac{1}{\sin 45^\circ \sin 46^\circ} + \frac{1}{\sin 47^\circ \sin 48^\circ} + \cdots + \frac{1}{\sin 133^\circ \sin 134^\circ} = \frac{1}{\sin n^\circ}.$$

(Source: AIME) **Hints:** 108, 86

3.74★ Find the sum of the values of x such that $\cos^3 3x + \cos^3 5x = 8\cos^3 4x \cos^3 x$, where x is measured in degrees and $100 < x < 200$. *(Source: AIME)* **Hints:** 178

3.75★ An angle x is chosen at random from the interval $0^\circ < x < 90^\circ$. Find the probability that the numbers $\sin^2 x$, $\cos^2 x$, and $\sin x \cos x$ are *not* the lengths of the sides of a triangle. *(Source: AIME)*

3.76★ Show that if $x + y + z = xyz$, then $\dfrac{2x}{1-x^2} + \dfrac{2y}{1-y^2} + \dfrac{2z}{1-z^2} = \dfrac{2x}{1-x^2} \cdot \dfrac{2y}{1-y^2} \cdot \dfrac{2z}{1-z^2}$. **Hints:** 153

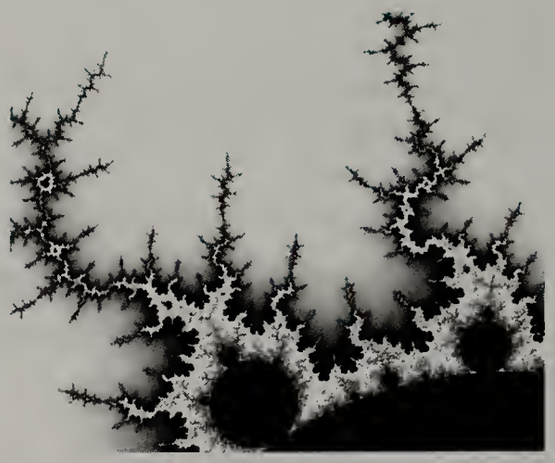
3.77★ The sequence $\{a_n\}$ satisfies $a_0 = 0$ and $a_{n+1} = \frac{8}{5}a_n + \frac{6}{5}\sqrt{4^n - a_n^2}$ for $n \geq 0$. Find a_{10}. *(Source: AIME)* **Hints:** 145, 216

3.78★ Prove that $\displaystyle\sum_{k=1}^{n} \arctan\frac{1}{2k^2} = \arctan\frac{n}{n+1}$.

Sidenote: ♪ On page 111, we asked what causes the oscillations we hear when tuning an instrument that is almost in tune, but not quite. To investigate, we graphed $y = \sin 30x + \sin 30x$, $y = \sin 31x + \sin 29x$, and $y = \sin 32x + \sin 28x$. Applying the sum-to-product identity gives us a clearer picture of what's going on in the graphs. While $\sin 30x + \sin 30x$ is just $2\sin 30x$, applying the sum-to-product identity for sine gives

$$\sin 31x + \sin 29x = 2\sin 30x \cos x,$$
$$\sin 32x + \sin 28x = 2\sin 30x \cos 2x.$$

In each case, the $2\sin 30x$ of the in-tune guitar is damped by an extra term that oscillates. Since $\cos 2x$ oscillates twice as fast (has half the period) as $\cos x$, the sound oscillates much faster for the $\sin 32x + \sin 28x$ guitar than it does for the $\sin 31x + \sin 29x$ guitar.



Geometry is the science of correct reasoning on incorrect figures. – George Polya

CHAPTER 4
Applications to Geometry

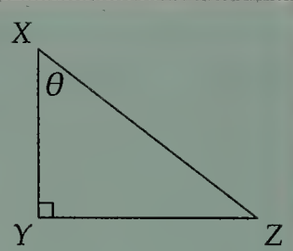
We explore many applications of trigonometric functions to geometry in this chapter. Throughout the chapter, you will encounter trig functions of angles that you cannot evaluate in your head or with pencil and paper, such as $\sin 14°$. When you encounter such expressions in this chapter, you can use a calculator to approximate them.

4.1 Right Triangle Trigonometry

Back on page 35, we proved the following:

Important:

If triangle XYZ is a right triangle with $\angle XYZ = 90°$ and $\angle X = \theta$, then



$$\sin\theta = \frac{YZ}{XZ}, \qquad \cos\theta = \frac{YX}{XZ}, \qquad \tan\theta = \frac{YZ}{XY}.$$

Another way of viewing this is: If θ is the measure of an acute angle in a right triangle, then we have

$$\sin\theta = \frac{\text{opposite leg}}{\text{hypotenuse}}, \qquad \cos\theta = \frac{\text{adjacent leg}}{\text{hypotenuse}}, \qquad \tan\theta = \frac{\text{opposite leg}}{\text{adjacent leg}}.$$

Some people remember these relationships with the mnemonic "**SOHCAHTOA**" (pronounced "so-cah-toe-ah"):

Sine: **O**pposite/**H**ypotenuse, **C**osine: **A**djacent/**H**ypotenuse, **T**angent: **O**pposite/**A**djacent.

In this section, we apply these relationships to a variety of problems.

Important: Throughout this chapter we will use some very common conventions when working with angles of a triangle and trigonometric functions of those angles. Specifically, when referring to the angles of $\triangle ABC$, we will often refer angles of the triangle solely by their vertices. That is, we will write $\angle A$, $\angle B$, and $\angle C$ rather than using three points to identify each angle. When writing trigonometric functions, we will often omit the $\angle$ symbol and simply write $\sin A$, which is the same as $\sin \angle BAC$. When referring to the side lengths of $\triangle ABC$, we often let $a = BC$, $b = AC$, and $c = AB$.

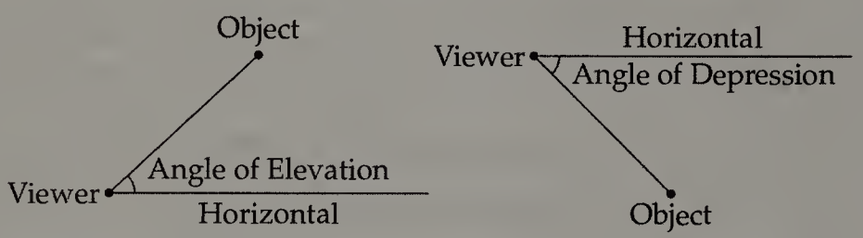
Problems

Problem 4.1:

(a) In $\triangle ABC$, we have $\angle ABC = 90°$, $AC = 6$, and $\angle BAC = 24°$. Find BC and AB.

(b) In $\triangle PQR$, we have $PQ = 7$, $QR = 24$, and $\angle PQR = 90°$. Find $\angle QPR$.

Problem 4.2: The **angle of elevation** is the angle above the horizontal at which a viewer must look to see an object that is higher than the viewer. Similarly, the **angle of depression** is the angle below the horizontal at which a viewer must look to see an object that is below the viewer.



(a) A surveyor measures the angle of elevation from her feet to the top of a building as 5°. The surveyor knows that the building is 500 feet tall. Assuming the ground is flat and level between the surveyor and the building, how far away is the surveyor from the building?

(b) A bee is on a hill looking at a building. The building is 400 feet tall. The angle of elevation from the bee to the top of the building is 4° and the angle of depression from the bee to the bottom of the building is 2°. What is the shortest distance the bee will have to fly to reach the building? (Answer to the nearest hundred feet.)

Problem 4.3: Suppose the slope of a line graphed in the Cartesian plane is m, where $m \neq 0$. Show that m equals the tangent of one of the angles that the line makes with the x-axis.

Problem 4.4: In right triangle ABC, $\angle ABC = 90°$ and $AC = 4$. If P is on $\overline{AC}$ such that $BP = 1$ and $\overline{BP} \perp \overline{AC}$, then find all possible values of $\angle BCP$. *(Source: CEMC)*

Problem 4.5: In $\triangle ABC$, let $a = BC$ and $b = AC$, and let $[ABC]$ be the area of $\triangle ABC$. Show that $[ABC] = \frac{1}{2}ab \sin C$.

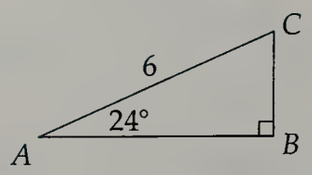
Problem 4.6: What is the largest possible area of a triangle that has two consecutive sides of length 10? Prove that no larger area is possible.

Problem 4.7: $ABCD$ is a square and M and N are the midpoints of $\overline{BC}$ and $\overline{CD}$, respectively. What is $\sin \angle MAN$? *(Source: AHSME)*

Problem 4.1:
(a) In $\triangle ABC$, we have $\angle ABC = 90°$, $AC = 6$, and $\angle BAC = 24°$. Find BC and AB.
(b) In $\triangle PQR$, we have $PQ = 7$, $QR = 24$, and $\angle PQR = 90°$. Find $\angle QPR$.

Solution for Problem 4.1:

(a) We have

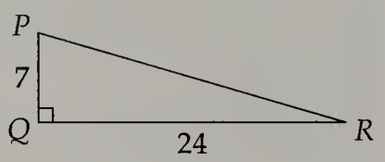


$$\sin \angle BAC = \frac{BC}{AC},$$
$$\cos \angle BAC = \frac{AB}{AC}.$$

Using a calculator to compute $\sin 24°$ and $\cos 24°$, we find

$$BC = AC \sin \angle BAC = 6 \sin 24° \approx 2.44,$$
$$AB = AC \cos \angle BAC = 6 \cos 24° \approx 5.48.$$

(b) We have



$$\tan \angle QPR = \frac{QR}{PQ} = \frac{24}{7}.$$

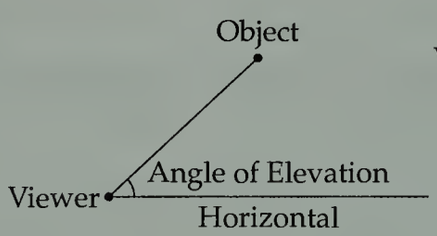
Using a calculator, we find that

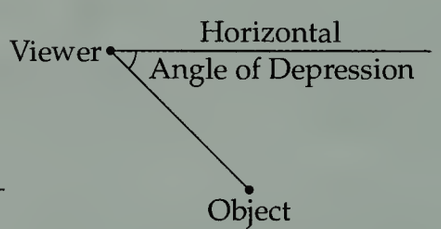
$$\angle QPR = \arctan \frac{24}{7} \approx 73.74°.$$

□

Trigonometry has a great many real-world applications. Surveyors and (perhaps more interesting to you) video game programmers use trigonometry frequently. Here's an example; we'll see several more throughout this chapter.

Problem 4.2: The **angle of elevation** is the angle above the horizontal at which a viewer must look to see an object that is higher than the viewer. Similarly, the **angle of depression** is the angle below the horizontal at which a viewer must look to see an object that is below the viewer.

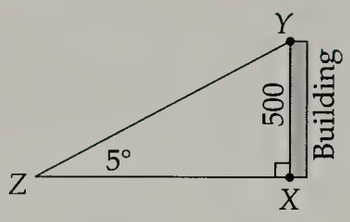




(a) A surveyor measures the angle of elevation from her feet to the top of a building as 5°. The surveyor knows that the building is 500 feet tall. Assuming the ground is flat and level between the surveyor and the building, how far away is the surveyor from the building?
(b) A bee is on a hill looking at a building. The building is 400 feet tall. The angle of elevation from the bee to the top of the building is 4° and the angle of depression from the bee to the bottom of the building is 2°. What is the shortest distance the bee will have to fly to reach the building? (Answer to the nearest hundred feet.)

Solution for Problem 4.2:

(a) The diagram at right depicts the building, with the surveyor at point Z. We seek ZX, which is one leg of $\triangle XYZ$, and we know the other leg of $\triangle XYZ$. When we apply trigonometric relationships to $\triangle XYZ$, we focus the side we know, the side we want, and the acute angle we know:



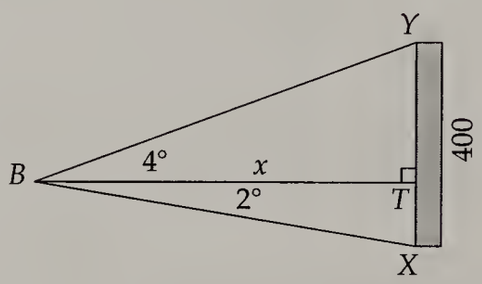
$$\tan \angle YZX = \frac{YX}{ZX}.$$

Solving for ZX gives us

$$ZX = \frac{YX}{\tan \angle YZX} = \frac{500}{\tan 5°} \approx 5715.$$

Therefore, the surveyor is approximately 5715 feet from the building.

(b) In the diagram, the bee is at point B, and the distance from the bee to the building is BT, where T is the foot of the perpendicular from the bee to the building. We have two right triangles, and we know the measures of the acute angles in those triangles, but we don't know any of the side lengths of those triangles. In hopes of finding useful relationships among lengths of segments in the diagram, we let $BT = x$, and try to find other lengths in the diagram in terms of x. We assign a variable to BT both because $\overline{BT}$ is a side of both right triangles, and because BT is the length we seek in the problem.



We know that $XY = 400$, so we focus on the segments along $\overline{XY}$. From $\triangle BTY$, we have $\tan \angle TBY = \frac{YT}{BT}$, so

$$YT = BT \tan \angle TBY = x \tan 4° \approx 0.070x.$$

From $\triangle BTX$, we have $\tan \angle TBX = \frac{XT}{BT}$, so

$$XT = BT \tan \angle TBX = x \tan 2° \approx 0.035x.$$

Since we know that $XY = 400$, we now have an equation for x. We have $XY = YT + XT \approx 0.105x$, so $0.105x \approx 400$, from which we find $x \approx 3800$. Therefore, the bee is approximately 3800 feet from the building.

Our key step in this solution was the first one:

Concept: Many geometry problems are solved by assigning variables to lengths or angles in the problems, and then finding other lengths or angles in terms of the variables. The goal then is to find an equation that you can solve.

□

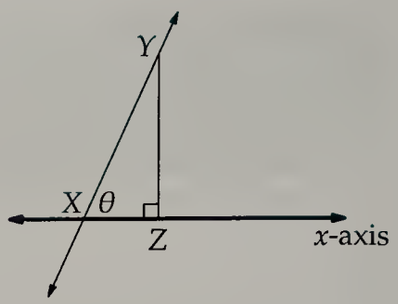
In addition to the many applications of trigonometry to Euclidean geometry, there are also applications to analytic geometry.

Problem 4.3: Suppose the slope of a line graphed in the Cartesian plane is m, where $m \neq 0$. Show that m equals the tangent of one of the angles that the line makes with the x-axis.

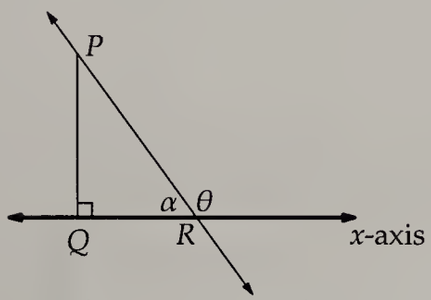
Extra! *There is a difference between not knowing and not knowing yet.*

– Shelia Tobias

Solution for Problem 4.3: We'll first address a line with positive slope. In the diagram at right, the line makes an acute angle of θ with the x-axis as shown. (We don't need the y-axis for this problem, so we haven't included it.) Point Y is on the line above the x-axis, and point Z is the foot of the altitude from Y to the x-axis. The slope of this line is positive, so we have $m = \frac{YZ}{XZ}$. From right triangle YXZ, we have



$$\tan\theta = \frac{YZ}{XZ} = m.$$



Turning to a line with negative slope, we start with the diagram at left. Point P is on the line above the x-axis, and Q is the foot of the altitude from P to the x-axis. As we move upward along the line, we go to the left, not the right, so we have $m = -\frac{PQ}{QR}$. However, in $\triangle PQR$, we have

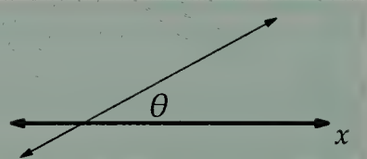
$$\tan\alpha = \frac{PQ}{QR} = -m.$$

Uh-oh. The problem says that the tangent of one of the angles the line makes with the x-axis should be equal to the slope, not the negative of the slope. Let's take a look at the obtuse angle the line makes with the x-axis. Since this angle together with α makes a straight angle, we have $\alpha + \theta = 180°$, so

$$\tan\theta = \tan(180° - \alpha) = \frac{\sin(180° - \alpha)}{\cos(180° - \alpha)} = \frac{\sin\alpha}{-\cos\alpha} = -\tan\alpha.$$

Therefore, we have $\tan\theta = -(-m) = m$, and once again, the tangent of one of the angles formed by the line and the x-axis equals the slope. □

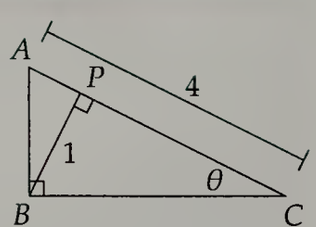
Important: A line with nonzero slope m must intersect the x-axis. The line and the x-axis form an angle θ that is to the right of the line and above the x-axis, as shown. Then, we have $\tan\theta = m$.



Our solution to Problem 4.3 gives us a quick way to visualize this relationship, so you shouldn't have to memorize it.

Problem 4.4: In right triangle ABC, $\angle ABC = 90°$ and $AC = 4$. If P is on $\overline{AC}$ such that $BP = 1$ and $\overline{BP} \perp \overline{AC}$, then find all possible values of $\angle BCP$. *(Source: CEMC)*

Solution for Problem 4.4: We let the desired angle be θ, and present two solutions:

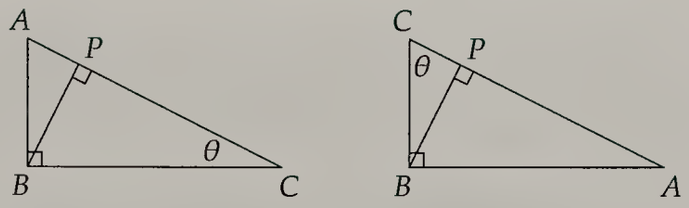


Solution 1: Find sides in terms of θ. We have one side length in right triangle ABC and one in right triangle BPC. If we could find one more side in either triangle, we could use the trigonometric relationships in a right triangle to find θ. We focus on $\overline{BC}$, since this segment is a side of both triangles. In $\triangle BPC$, we have $\sin\theta = \frac{BP}{BC} = \frac{1}{BC}$, so $BC = \frac{1}{\sin\theta}$. In $\triangle ABC$, we have $\cos\theta = \frac{BC}{AC} = \frac{BC}{4}$, so $BC = 4\cos\theta$. (Notice that in each case, we investigate the trig function that we can express in terms of BC and a side length that we know.) We now have two different expressions for BC, so these expressions must be equal:

$$\frac{1}{\sin\theta} = 4\cos\theta.$$

Multiplying both sides by $\sin\theta$ gives $4\sin\theta\cos\theta = 1$. Applying the sine double angle identity (in "reverse") makes our equation $2\sin 2\theta = 1$, so $\sin 2\theta = \frac{1}{2}$. Since θ is acute, there are two possible solutions, $2\theta = 30°$ or $2\theta = 150°$. The former gives $\theta = 15°$ and the latter gives $\theta = 75°$. So, the two possible values of $\angle BCP$ are $15°$ and $75°$.

We can see why the possibilities for $\angle C$ are complementary in the diagrams at right. Specifically, for any triangle ABC that satisfies the problem, we can swap the labels A and C to produce another triangle that satisfies the problem. Since $\angle A + \angle C = 90°$, we know that these two possibilities for $\angle C$ must be complementary.



Solution 2: Area. All those right angles suggest that we might be able to use area to solve this problem. Since $\overline{BP}$ is an altitude of $\triangle ABC$ to side $\overline{AC}$, we have

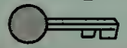
$$[ABC] = \frac{(BP)(AC)}{2} = 2,$$

where $[ABC]$ denotes the area of $\triangle ABC$. We can also express the area of $\triangle ABC$ as half the product of its legs. Using trigonometric relationships in $\triangle ABC$, we can express both legs in terms of θ. As explained above, we have $BC = 4\cos\theta$. We also have $\sin\theta = \frac{AB}{AC} = \frac{AB}{4}$, so $AB = 4\sin\theta$. Therefore,

$$[ABC] = \frac{(AB)(BC)}{2} = 8\sin\theta\cos\theta.$$

We now have two expressions for $[ABC]$. Setting these equal gives $8\sin\theta\cos\theta = 2$, and the rest of the solution proceeds as before. □

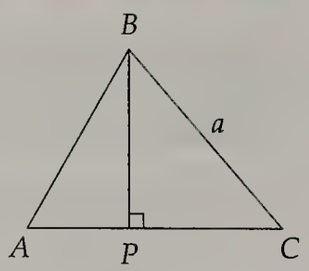
Both solutions involve the key first step of assigning a variable and expressing lengths in terms of that variable. The second offers a very useful general geometric problem solving tactic:

Concept:  Area can be a surprisingly powerful problem solving tool, even in problems that don't appear to be about area.

Having seen one example of a relationship between trigonometric functions and area, let's now take a look at a more general one.

Problem 4.5: In $\triangle ABC$, let $a = BC$, $b = AC$, and $c = AB$, and let $[ABC]$ be the area of $\triangle ABC$. Show that $[ABC] = \frac{1}{2}ab\sin C$.

Solution for Problem 4.5: Our usual formula for triangle area is "one-half base times height." We seek a way to express a height of the triangle in terms of $\angle C$, hoping to convert the "one-half base times height" formula to the desired formula. So, we draw altitude $\overline{BP}$, and we have



$$[ABC] = \frac{(BP)(AC)}{2} = \frac{(BP)(b)}{2}.$$

What's wrong with this finish:

Bogus Solution: The altitude introduces a right triangle with acute angle $\angle C$. We have $\sin C = \frac{BP}{BC} = \frac{BP}{a}$, so $BP = a\sin C$, and

$$[ABC] = \frac{(BP)(b)}{2} = \frac{1}{2}ab\sin C.$$

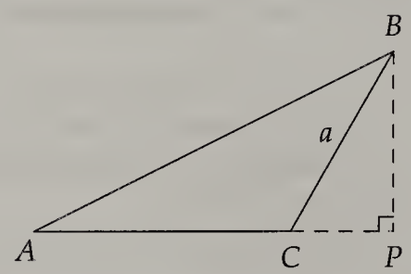
So, for any $\triangle ABC$ with $a = BC$ and $b = AC$, we have $[ABC] = \frac{1}{2}ab\sin C$.

The problem here is that our proof does not adequately cover every possible triangle ABC. The Bogus Solution only covers the the case in which $\angle C$ is acute. What happens if we have $\angle C = 90°$? Then P is at C, which presents some problems. Worse yet, what if $\angle C > 90°$? We still have to address these two cases. (Our proof above does handle the cases in which $\angle A = 90°$ or $\angle A > 90°$; make sure you see why.)

WARNING!! When proving a statement for all possible triangles, make sure your proof is valid for all possible triangles. Specifically, you may have to address right and obtuse triangles as separate cases from acute triangles.

Fortunately, taking care of the other cases is pretty easy. If $\angle C = 90°$, then we have $[ABC] = \frac{(AC)(BC)}{2} = \frac{1}{2}ab$. Since $\sin 90° = 1$, the proposed formula also gives $[ABC] = \frac{1}{2}ab\sin 90° = \frac{1}{2}ab$, so the proposed formula gives the correct area when $\angle C = 90°$.

If $\angle C > 90°$, we have the diagram at right, where again our first step is drawing the altitude from B. From $\triangle BCP$, we have $\sin\angle BCP = \frac{BP}{BC}$, so $BP = BC\sin\angle BCP = a\sin\angle BCP$. Therefore, we have



$$[ABC] = \frac{(AC)(BP)}{2} = \frac{(b)(a\sin\angle BCP)}{2} = \frac{1}{2}ab\sin\angle BCP.$$

That's close, but not quite the same as our desired formula, $[ABC] = \frac{1}{2}ab\sin\angle ACB$. However, we have $\angle BCP = 180° - \angle ACB$, so

$$[ABC] = \frac{1}{2}ab\sin\angle BCP = \frac{1}{2}ab\sin(180° - \angle ACB).$$

Applying the identity $\sin\theta = \sin(180° - \theta)$ gives $[ABC] = \frac{1}{2}ab\sin(180° - \angle ACB) = \frac{1}{2}ab\sin\angle ACB$, so the desired formula works when $\angle C$ is obtuse as well.

Important: If $a = BC$, $b = AC$, and $[ABC]$ is the area of $\triangle ABC$, then

$$[ABC] = \frac{1}{2}ab\sin C.$$

□

Problem 4.6: What is the largest possible area of a triangle that has two consecutive sides of length 10? Prove that no larger area is possible.

Solution for Problem 4.6: Let the triangle be $\triangle XYZ$ with $XY = YZ = 10$. We can relate the area of XYZ to the two sides with the formula we just proved:

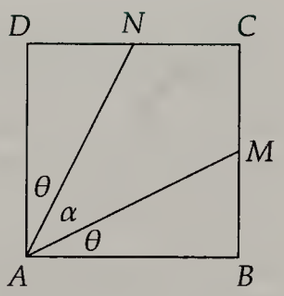
$$[XYZ] = \frac{1}{2}(XY)(YZ)\sin\angle XYZ = 50\sin\angle XYZ.$$

The maximum possible value of $\sin\angle XYZ$ is 1, which occurs when $\angle XYZ = 90°$. Therefore, the maximum possible area of $\triangle XYZ$ is 50, which occurs when $\angle XYZ$ is a right angle. □

Problem 4.7: $ABCD$ is a square and M and N are the midpoints of $\overline{BC}$ and $\overline{CD}$, respectively. What is $\sin\angle MAN$? *(Source: AHSME)*

Solution for Problem 4.7: We present two solutions.

Solution 1: Trig Identities. We start with the diagram at right, in which α is the angle whose sine we seek. While we don't have a right triangle that includes α as an acute angle, we can easily relate α to acute angles $\angle BAM$ and $\angle DAN$ in right triangles BAM and DAN. Because $\triangle BAM$ and $\triangle DAN$ are congruent by SAS Congruence ($AB = AD$, $\angle B = \angle D$, and $BM = DN$), we have $\angle DAN = \angle BAM$. Letting both of these angles have measure θ, we have



$$\sin\alpha = \sin(90^\circ - 2\theta) = \cos 2\theta = 2\cos^2\theta - 1.$$

(We used the two identities $\sin(90^\circ - x) = \cos x$ and $\cos 2x = 2\cos^2 x - 1$ here. We could have used either of the other cosine double angle identities, as well.)

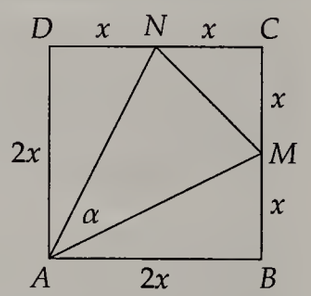
From right triangle MAB, we have $\cos\theta = \frac{AB}{AM}$. Since M is the midpoint of $\overline{BC}$ and $ABCD$ is a square, we have $MB = \frac{BC}{2} = \frac{AB}{2}$. Therefore, applying the Pythagorean Theorem to $\triangle BAM$ gives

$$AM = \sqrt{AB^2 + BM^2} = \sqrt{AB^2 + \frac{AB^2}{4}} = \frac{\sqrt{5}}{2}AB,$$

which means $\frac{AB}{AM} = \frac{2}{\sqrt{5}}$ and we have

$$\sin\alpha = 2\cos^2\theta - 1 = 2\left(\frac{AB}{AM}\right)^2 - 1 = 2\left(\frac{2}{\sqrt{5}}\right)^2 - 1 = \frac{3}{5}.$$

Solution 2: Area. We let $AB = AD = 2x$, so $MB = MC = NC = ND = x$. (We let $AB = 2x$ rather than $AB = x$ to avoid introducing fractions when computing the shorter segments.) We draw $\overline{MN}$ to divide the square into three right triangles and $\triangle AMN$. The square has area $4x^2$, and we have $[AMB] = [AND] = x^2$ and $[CNM] = \frac{x^2}{2}$. We therefore have $[AMN] = 4x^2 - x^2 - x^2 - \frac{x^2}{2} = \frac{3x^2}{2}$. We have $AN = AM = x\sqrt{5}$ from the Pythagorean Theorem, so



$$[AMN] = \frac{1}{2}(AN)(AM)\sin\alpha = \frac{5x^2}{2}\sin\alpha.$$

Therefore, we must have $\frac{5x^2}{2}\sin\alpha = \frac{3x^2}{2}$, so $\sin\alpha = \frac{3}{5}$. Once again, we have used area to tackle a problem that doesn't appear to be about area.

Notice that our answer doesn't depend on x. Because scaling the diagram up or down doesn't affect $\angle MAN$, we could have chosen a specific value for AB rather than letting it be x. So, we could have simply let $AB = AD = 2$ and proceeded as above to find $\sin\alpha$. □

Sidenote: In our solutions to Problem 4.7, we discovered an angle with a sine of 3/5. One of the acute angles of a 3-4-5 triangle has sine equal to 3/5. Perhaps we could solve the problem by finding a 3-4-5 triangle with $\angle MAN$ as one of its acute angles. See if you can find it with a little "outside-the-box" thinking. *Answer on page 153.*

Exercises

4.1.1 If $\angle PQR = 90^\circ$, $\angle QRP = 42^\circ$, and $QP = 8$, then what are PR and QR to the nearest hundredth?

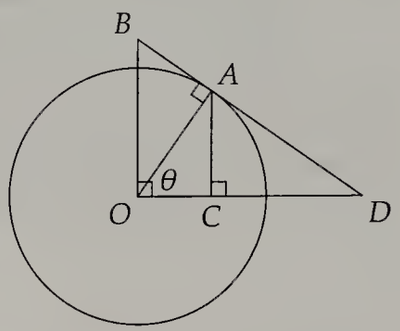
4.1.2 If $XY = 4YZ$ and $\angle XYZ = 90^\circ$, then what is $\angle XZY$ to the nearest degree?

4.1.3 Find the measure of an acute angle formed by the x-axis and the graph of $2x + 3y = 7$. (Answer to the nearest degree.)

4.1.4 Find the area of $\triangle ABC$ if $AB = 4BC = 12$ and $\angle ABC = 30°$.

4.1.5 I'm standing at the peak of a mountain that is 14,000 feet above sea level. The angle of depression from this peak to a nearby smaller peak is 4°. On my map, these two peaks are represented by points that are 1 inch apart. If each inch on my map represents 1.2 miles, and there are 5280 feet in a mile, then how many feet above sea level is the second peak?

4.1.6 The radius of the circle at right is 1. For each of the following trigonometric expressions, find a segment in the diagram that has length equal to the trigonometric expression: $\sin\theta$, $\cos\theta$, $\sec\theta$, $\csc\theta$, $\tan\theta$, $\cot\theta$. (Note: You are *not* asked to express each trig function in terms of multiple segments in the diagram; you must find a segment whose length *equals* the corresponding trig function.)



4.1.7 A, B, and C are vertices of a cube such that $\overline{AB}$ is an interior diagonal of the cube and $\overline{AC}$ is a diagonal of a face of the cube. Find $\cos\angle ABC$.

4.1.8★ Square $ABCD$ has center O and $AB = 900$. Points E and F are on $\overline{AB}$ with $AE < BF$ and E between A and F such that $\angle EOF = 45°$ and $EF = 400$. Find BF. *(Source: AIME)*

4.2 Law of Cosines

Problems

Problem 4.8: In $\triangle ABC$, let $AC = 14$, $BC = 12$, and $\angle C = 34°$. In this problem, we find AB to the nearest hundredth.

(a) Draw altitude $\overline{BX}$ from B to $\overline{AC}$. Find CX and BX to the nearest hundredth.

(b) Find XA to the nearest hundredth.

(c) Use parts (a) and (b) to find AB to the nearest hundredth.

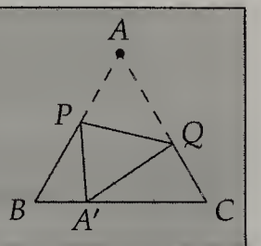
Problem 4.9: Let $\triangle ABC$ be an acute triangle with $a = BC$, $b = AC$, and $c = AB$. Use Problem 4.8 as a guide to prove that

$$c^2 = a^2 + b^2 - 2ab\cos C.$$

Problem 4.10: An airplane leaves an aircraft carrier and flies due south at 400 km/h. The carrier proceeds 60° east of north at 32 km/h. If the plane has enough fuel for 5 hours of flying, what is the maximum distance south the pilot can travel, so that the fuel remaining will allow a safe return to the carrier? (You may assume Earth is flat in this problem.) *(Source: CEMC)*

Problem 4.11: In $\triangle ABC$, we have $AB = 5$, $BC = 7$, and $AC = 8$. Find $\angle BAC$.

Problem 4.12: Equilateral triangle ABC has been creased and folded so that vertex A now rests at A' on $\overline{BC}$ as shown. If $BA' = 1$ and $A'C = 2$, then what is the length of crease $\overline{PQ}$? *(Source: AMC 12)*



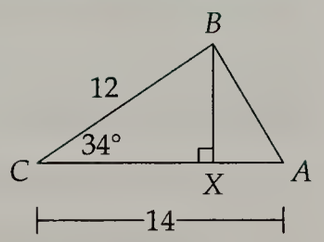
(a) Why do 60° and 120° angles suggest trying the Law of Cosines?

(b) Find PQ. **Hints:** 228

Problem 4.13: In quadrilateral $ABCD$, we have $\angle A = \angle C$, $AB = CD = 180$, and $AD \neq BC$. The perimeter of $ABCD$ is 640. Find $\cos A$. *(Source: AIME)* **Hints:** 175

Problem 4.8: In $\triangle ABC$, let $AC = 14$, $BC = 12$, and $\angle C = 34°$. Find AB to the nearest hundredth.

Solution for Problem 4.8: We know how to use trigonometry to find side lengths in right triangles, so we start by drawing an altitude from B to $\overline{AC}$ as shown. This creates right triangle $\triangle BCX$ with the 34° angle as one of its acute angles. From right triangle $\triangle CBX$, we have



$$\frac{BX}{BC} = \sin C \approx 0.559,$$

so $BX \approx 0.559(BC) \approx 6.71$. Similarly, we have $CX/BC = \cos C \approx 0.829$, so $CX \approx 9.95$. This doesn't tell us AB yet, but we now have the length of one leg of $\triangle BXA$. If we can find the other, we can use the Pythagorean Theorem to find AB. Fortunately, XA is easy to find: $XA = AC - CX \approx 4.05$. Now, we can use the Pythagorean Theorem to find

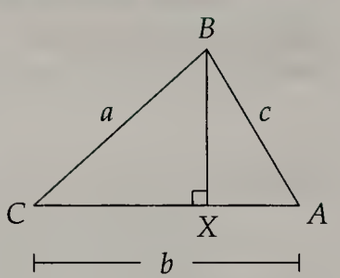
$$AB = \sqrt{BX^2 + XA^2} \approx 7.84.$$

□

In Problem 4.8, we were given two side lengths of a triangle and the measure of the angle between these two sides. We then found the third side. There was nothing particularly special about the side lengths or the angle. We might be able to follow essentially the same process for any triangle. Let's give it a try.

Problem 4.9: Let $\triangle ABC$ be an acute triangle with $a = BC$, $b = AC$, and $c = AB$. Find a formula for c in terms of a, b, and $\angle C$.

Solution for Problem 4.9: We use Problem 4.8 as a guide. In fact, this problem is essentially the same as Problem 4.8, but with variables a, b, and $\angle C$ in place of the numbers that were given in the earlier problem. We can use the same steps. We draw altitude $\overline{BX}$ from B to $\overline{AC}$. Then, we have $\sin C = BX/BC$, so $BX = BC \sin C = a \sin C$. We also have $\cos C = CX/BC$, so $CX = BC \cos C = a \cos C$. Therefore, we have $AX = AC - CX = b - a \cos C$. Next we apply the Pythagorean Theorem to $\triangle ABX$ to find $AB^2 = BX^2 + AX^2$. Substituting our expressions for these three sides gives us



$$\begin{aligned} c^2 &= a^2 \sin^2 C + (b - a \cos C)^2 \\ &= a^2 \sin^2 C + b^2 - 2ab \cos C + a^2 \cos^2 C \\ &= a^2(\sin^2 C + \cos^2 C) + b^2 - 2ab \cos C. \end{aligned}$$

Since $\sin^2 C + \cos^2 C = 1$, we have

$$c^2 = a^2 + b^2 - 2ab \cos C.$$

□

We can test our new formula with the data from Problem 4.8.

Concept: Once you think you've found a new formula, check your work by trying the formula on specific examples you have solved without the formula.

In Problem 4.8, we have $a = 12$, $b = 14$, and $\angle C = 34°$, so we have

$$c^2 = a^2 + b^2 - 2ab \cos C \approx 61.44.$$

Taking the square root of both sides gives $c \approx 7.84$, which agrees with our answer from Problem 4.8.

With a little more casework (which you'll supply as an Exercise), we can show that this equation holds for any triangle ABC.

Important: Let $a = BC$, $b = AC$, and $c = AB$ in $\triangle ABC$. The **Law of Cosines** states that

$$c^2 = a^2 + b^2 - 2ab\cos C.$$

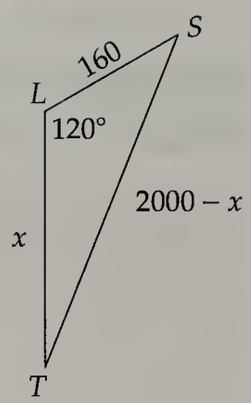
Notice that when $\angle C = 90°$, we have $\cos C = 0$, so the Law of Cosines becomes $c^2 = a^2 + b^2$, which is just the Pythagorean Theorem.

Problem 4.8 is just a specific example of Problem 4.9, so we call Problem 4.9 a **generalization** of Problem 4.8.

Concept: Solving a specific example of a general problem can often provide a guide for solving the general problem.

Problem 4.10: An airplane leaves an aircraft carrier and flies due south at 400 km/h. The carrier proceeds 60° east of north at 32 km/h. If the plane has enough fuel for 5 hours of flying, what is the maximum distance south the pilot can travel, so that the fuel remaining will allow a safe return to the carrier? (You may assume Earth is flat in this problem.) *(Source: CEMC)*

Solution for Problem 4.10: We start with a diagram, including the path of the ship and the path of the plane. The plane leaves the ship at L, flies south to T before turning, and then flies to S, where it lands on the ship. Because the plane flies south and the ship goes 60° east of north, we have $\angle TLS = 180° - 60° = 120°$. If the plane flies for 5 hours, then the ship moves $32 \cdot 5 = 160$ km and the plane flies $400 \cdot 5 = 2000$ km. Therefore, $LS = 160$ and $LT + TS = 2000$. We let the desired distance, LT, be x, so $TS = 2000 - x$, as shown. We have expressions for all three sides of $\triangle LST$, and we know one angle, so we apply the Law of Cosines:



$$TS^2 = LT^2 + LS^2 - 2(LT)(LS)\cos\angle TLS.$$

Inserting our expressions and values for lengths and the angle, we have

$$(2000 - x)^2 = x^2 + 160^2 - 2(x)(160)\cos 120°.$$

We have $\cos 120° = -\frac{1}{2}$ (which is one reason the Law of Cosines is often helpful in problems involving 120° angles), and expanding the left side of the equation gives

$$2000^2 - 4000x + x^2 = x^2 + 160^2 + 160x.$$

Solving this equation gives $x \approx 955$, so the plane can fly approximately 955 km south. □

Problem 4.11: In $\triangle ABC$, we have $AB = 5$, $BC = 7$, and $AC = 8$. Find $\angle BAC$.

Solution for Problem 4.11: We are given all three side lengths of the triangle, so we have all the information in the Law of Cosines except the angle measure. Therefore, we can solve for $\cos\angle BAC$ with the Law of Cosines. We have

$$BC^2 = AB^2 + AC^2 - 2(AB)(AC)\cos\angle BAC,$$

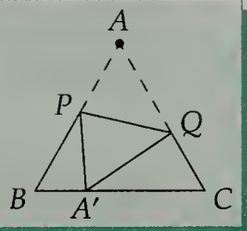
so

$$\cos\angle BAC = \frac{AB^2 + AC^2 - BC^2}{2(AB)(AC)} = \frac{25 + 64 - 49}{2(5)(8)} = \frac{1}{2}.$$

Since $0° < \angle BAC < 180°$, $\cos\angle BAC = \frac{1}{2}$ gives us $\angle BAC = 60°$. □

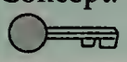
Concept: The Law of Cosines can be used to find angle measures as well as side lengths.

Problem 4.12: Equilateral triangle ABC has been creased and folded so that vertex A now rests at A' on $\overline{BC}$ as shown. If $BA' = 1$ and $A'C = 2$, then what is the length of crease $\overline{PQ}$? *(Source: AMC 12)*



Solution for Problem 4.12: We know that the side length of the original equilateral triangle is $BA' + A'C = 3$, but we can't immediately find any other lengths. So, we assign a variable and hope to express other side lengths in terms of that variable. We might start by letting $z = PQ$, but we don't see any way to express another side length in terms of z. Therefore, we start by letting $x = BP$, which gives us $AP = AB - BP = 3 - x$. But now we seem stuck.

Concept: When stuck on a problem, focus on information you haven't used yet.



We haven't used the fact that $\triangle PA'Q$ is the reflection of $\triangle PAQ$ over $\overline{PQ}$. This tells us that $PA' = PA$, so $PA' = 3 - x$. Now, we have expressions for the lengths of all three sides of $\triangle PA'B$. We also know that $\angle PBA' = 60°$, so we apply the Law of Cosines:

$$(PA')^2 = BP^2 + (BA')^2 - 2(BP)(BA')\cos\angle PBA',$$

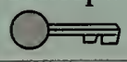
so

$$(3-x)^2 = x^2 + 1 - 2x\cos 60°.$$

Since $\cos 60° = \frac{1}{2}$, we have $(3-x)^2 = x^2 - x + 1$. Expanding the left side gives $9 - 6x + x^2 = x^2 - x + 1$, from which we find $x = \frac{8}{5}$. Therefore, we have $BP = \frac{8}{5}$ and $PA' = PA = 3 - BP = \frac{7}{5}$.

We can do the same thing with $\overline{AC}$!

Concept: If a tactic gives some information in a problem, but doesn't completely solve it, then try applying that tactic to the problem again in another way.



We let $QC = y$, so $A'Q = AQ = 3 - y$. Applying the Law of Cosines to $\triangle A'QC$ gives

$$(A'Q)^2 = (A'C)^2 + QC^2 - 2(A'C)(QC)\cos\angle A'CQ,$$

so

$$(3-y)^2 = 4 + y^2 - 2(2)(y)\cos 60°.$$

Solving this equation for y gives us $y = \frac{5}{4}$, so $QC = \frac{5}{4}$ and $A'Q = AQ = 3 - QC = \frac{7}{4}$.

One more time! We have $AP = \frac{7}{5}$, $AQ = \frac{7}{4}$, and $\angle PAQ = 60°$, so we apply the Law of Cosines to find

$$\begin{aligned} PQ^2 = AP^2 + AQ^2 - 2(AP)(AQ)\cos\angle PAQ &= \left(\frac{7}{5}\right)^2 + \left(\frac{7}{4}\right)^2 - 2\left(\frac{7}{5}\right)\left(\frac{7}{4}\right)\left(\frac{1}{2}\right) \\ &= \frac{49}{25} + \frac{49}{16} - \frac{49}{20} = \frac{49\cdot 16 + 49\cdot 25 - 49\cdot 20}{400} = \frac{49\cdot 21}{400}. \end{aligned}$$

Taking the square root gives $PQ = 7\sqrt{21}/20$. □

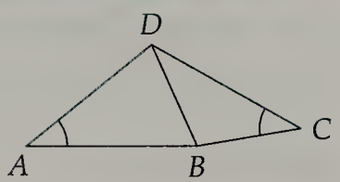
Each of the previous three problems featured a 60° or 120° angle. Usually, when we see 60° or 120°, our first strategy is to look for equilateral triangles or 30-60-90 triangles. But since $\cos 60° = \frac{1}{2}$ and $\cos 120° = -\frac{1}{2}$, the Law of Cosines has a particularly simple form when the angle is 60° or 120°.

Concept: If a geometry problem contains a 60° or 120° angle, you may be able to tackle it with the Law of Cosines.

WARNING!! The Law of Cosines can lead to some pretty heavy algebra, so usually we look for equilateral triangles or 30-60-90 triangles before pulling out the Law of Cosines on a problem that involves 60° or 120° angles.

Problem 4.13: In quadrilateral $ABCD$, we have $\angle A = \angle C$, $AB = CD = 180$, and $AD \neq BC$. The perimeter of $ABCD$ is 640. Find $\cos A$. *(Source: AIME)*

Solution for Problem 4.13: That the question asks for the cosine of an angle suggests that we either build right triangles or we apply the Law of Cosines. But we don't have any triangles. So, we make some. The obvious candidate for building triangles is to draw diagonals. We don't want to draw diagonal $\overline{AC}$, since that will break up two angles that we know are equal, and we'll probably have to use that angle equality. So, we try drawing $\overline{BD}$ first, thereby creating two triangles that share a side and have equal angles. We apply the Law of Cosines to each, using $\angle A$ in $\triangle ABD$ and $\angle C$ in $\triangle CBD$:



$$BD^2 = AB^2 + AD^2 - 2(AB)(AD)\cos A,$$
$$BD^2 = CB^2 + CD^2 - 2(CB)(CD)\cos C.$$

Setting our two expressions for BD^2 equal, and noting that $\cos C = \cos A$ because $\angle A = \angle C$, gives us

$$AB^2 + AD^2 - 2(AB)(AD)\cos A = CB^2 + CD^2 - 2(CB)(CD)\cos A.$$

We also have $AB = CD = 180$, so $AB^2 = CD^2$, and the equation is now

$$AD^2 - 2(180)(AD)\cos A = CB^2 - 2(180)(CB)\cos A.$$

To solve for $\cos A$, we group the $\cos A$ terms on the left, and we have

$$2(180)(CB)\cos A - 2(180)(AD)\cos A = CB^2 - AD^2,$$

so

$$\cos A = \frac{CB^2 - AD^2}{2(180)(CB - AD)}.$$

(This step only avoids dividing by 0 because we are told that $AD \neq BC$. This gives us some confidence that we're on the right track.) Factoring the numerator of the right side as a difference of squares, we have

$$\cos A = \frac{(CB - AD)(CB + AD)}{2(180)(CB - AD)} = \frac{CB + AD}{2(180)}.$$

Now, we just have to find $CB + AD$. We go back to the problem, and look for information we haven't used yet. We haven't used the perimeter. That involves the sum of sides! Let's try it. We have $AB + CD + CB + AD = 640$ and $AB = CD = 180$, so $CB + AD = 640 - 180 - 180 = 280$, and we have

$$\cos A = \frac{CB + AD}{2(180)} = \frac{280}{2(180)} = \frac{280}{360} = \frac{7}{9}.$$

□

Exercises

4.2.1 Complete our proof of the Law of Cosines by proving it for obtuse and right triangles.

4.2.2 Two airplanes take off from the same airport at the same time. One flies due west at 200 miles per hour. The other flies 40° north of due east at 250 miles per hour. To the nearest mile, how many miles apart are the planes 90 minutes after takeoff?

4.2.3 Prove each of the following:

(a) If $AB^2 = BC^2 + CA^2$, then $\angle ACB$ is right.

(b) If $AB^2 > BC^2 + CA^2$, then $\angle ACB$ is obtuse.

(c) If $AB^2 < BC^2 + CA^2$, then $\angle ACB$ is acute.

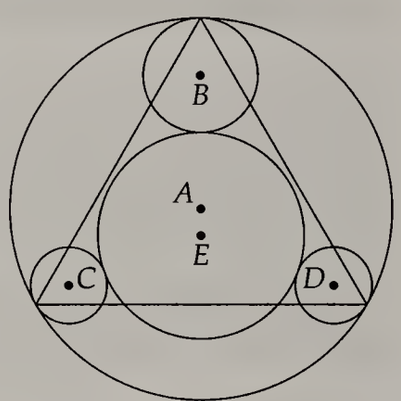
In each case, is the converse of the given statement true? Why or why not? (The **converse** of a statement of the form "If X, then Y" is "If Y, then X.")

4.2.4 One side of a triangle has length twice that of another side, and the third side has length 6. If one angle of the triangle is 120°, then what are the possible values of the lengths of the sides of the triangle?

4.2.5 If $XY = 3$, $YZ = 5$, and $ZX = 7$, then what is $\angle XYZ$?

4.2.6 Quadrilateral $ABCD$ is inscribed in a circle. If $AB = 2$, $BC = 3$, $CD = 4$, and $DA = 6$, then what is AC? *(Source: CEMC)*

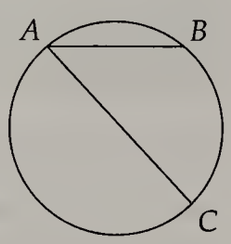
4.2.7⋆ Equilateral triangle $\mathcal{T}$ is inscribed in circle A, which has radius 10. Circle B with radius 3 is internally tangent to circle A at one vertex of $\mathcal{T}$. Circles C and D, both with radius 2, are internally tangent to circle A at the other two vertices of $\mathcal{T}$. Circles B, C, and D are all externally tangent to circle E. Find the radius of circle E. *(Source: AIME)*



4.2.8⋆ Triangle ABC has a right angle at B and contains a point P for which $PA = 10$, $PB = 6$, and $\angle APB = \angle BPC = \angle CPA$. Find PC. *(Source: AIME)* **Hints:** 205

4.3 Law of Sines

We start with an important fact from elementary geometry that will be helpful in this section. In the diagram at right, we say that $\angle BAC$ is **inscribed** in arc $\overset{\frown}{BC}$ of the circle because A, B, and C are all on the circle. The measure of an inscribed angle equals half the measure of the arc it intercepts, so $\angle A = \frac{\overset{\frown}{BC}}{2}$. If you are not familiar with this relationship, try proving it yourself before looking it up in a geometry text or online. (This relationship, and many others involving angles and circles, is covered in Art of Problem Solving's *Introduction to Geometry*.)



Problems

Problem 4.14: In $\triangle PQR$, we have $PR = 12$, $\angle QPR = 66°$, and $\angle PRQ = 63°$. In this problem, we find PQ and QR.

(a) Let T be the foot of the altitude from P to $\overline{QR}$. Find PT to the nearest hundredth.

(b) Use PT to find PQ to the nearest hundredth.

(c) Find QR to the nearest hundredth.

Problem 4.15: Suppose that $\triangle ABC$ is an acute triangle with $a = BC$, $b = AC$, and $c = AB$. Prove that

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}.$$

Problem 4.16: Let points P and Q be points on a shoreline that are 1.5 miles apart. Let B be the base of a radio tower on a distant island, and T be the top of the tower. A surveyor uses an angle measurement tool to determine that $\angle PQB = 58°$ and $\angle QPB = 72°$. She also measures that the angle of elevation from P to the top of the tower is $1°$. In this problem, we find the height of the tower in feet.

(a) Find PB.

(b) How tall is the tower in feet?

Problem 4.17: In this problem, we prove the **Extended Law of Sines**, which states that if $a = BC$, $b = AC$, $c = AB$, and R is the radius of circumcircle of $\triangle ABC$, then

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C} = 2R.$$

(Note: The circumcircle of a triangle is the circle that passes through all three vertices of the triangle.)

(a) Start with acute triangle $\triangle ABC$. Let D be the point on the circumcircle such that $\overline{AD}$ is a diameter of the circle. What angle of $\triangle ABC$ equals $\angle ADC$, and why must these angles be equal? **Hints:** 229

(b) Show that $\dfrac{b}{\sin \angle ABC} = 2R$, where R is the circumradius of $\triangle ABC$.

(c) Prove that the Extended Law of Sines holds for right and obtuse triangles.

Problem 4.18:

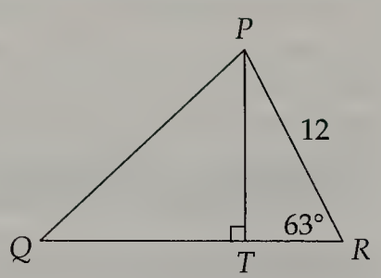
(a) Let $\triangle ABC$ and $\triangle DEF$ be triangles such that $AB = DE = 7$, $BC = EF = 5$, and $\angle CAB = \angle FDE = 30°$. Must $\sin \angle ACB = \sin \angle DFE$?

(b) Must $\triangle ABC$ and $\triangle DEF$ be congruent? If not, can we deduce anything about the relationship between $\angle ACB$ and $\angle DFE$?

Problem 4.19: In triangle ABC, $AB = 13$, $BC = 15$, and $CA = 14$. Point D is on $\overline{BC}$ with $CD = 6$. Point E is on $\overline{BC}$ such that $CE > CD$ and $\angle BAE = \angle CAD$. Find BE. *(Source: AIME)* **Hints:** 241, 179

Problem 4.14: Suppose $PR = 12$, $\angle QPR = 66°$, and $\angle PRQ = 63°$ in $\triangle PQR$. Find PQ and RQ to the nearest hundredth.

Solution for Problem 4.14: We start by building a right triangle, since we know how to use trigonometry to find lengths of sides in a right triangle. We draw altitude $\overline{PT}$ to side $\overline{QR}$. From right triangle $\triangle PTR$, we have $PT/PR = \sin R$, so $PT = PR \sin R \approx 10.69$. We also have $TR/PR = \cos R$, so $TR = PR \cos R \approx 5.45$.



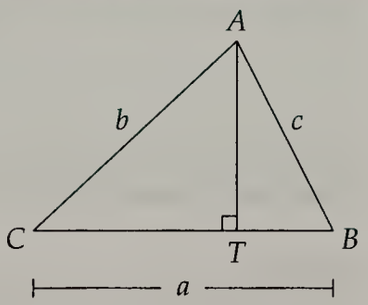
Now, we can use right triangle $\triangle PQT$ to find lengths QT and PQ. First, we find that $\angle Q = 180° - 66° - 63° = 51°$. We have $PT/PQ = \sin Q$, so we have $PQ = PT/(\sin Q) \approx 13.76$. We also have $PT/QT = \tan Q$, so $QT = PT/(\tan Q) \approx 8.66$. Finally, we have $QR = QT + TR \approx 14.11$. □

Let's try using our approach in Problem 4.14 to produce another "law."

Problem 4.15: Suppose that $\triangle ABC$ is an acute triangle with $a = BC$, $b = AC$, and $c = AB$. Prove that

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}.$$

Solution for Problem 4.15: We use Problem 4.14 as a guide. We draw altitude $\overline{AT}$ to $\overline{BC}$. From right triangle $\triangle ATC$, we have $\sin C = AT/AC$, so $AT = AC \sin C = b \sin C$. From right triangle $\triangle ABT$, we have $\sin B = AT/AB$, so $AT = AB \sin B = c \sin B$. These two expressions for AT must be equal, so we have $b \sin C = c \sin B$. Dividing this equation by $\sin B$ and by $\sin C$, we have



$$\frac{b}{\sin B} = \frac{c}{\sin C}.$$

We can follow essentially the same steps starting with the altitude from C to $\overline{AB}$ to show that $b/(\sin B) = a/(\sin A)$. So, we have

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}.$$

□

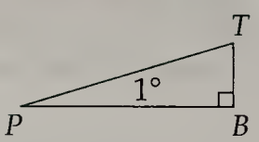
Important: Let $a = BC$, $b = AC$, and $c = AB$ in $\triangle ABC$. The **Law of Sines** states that

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}.$$

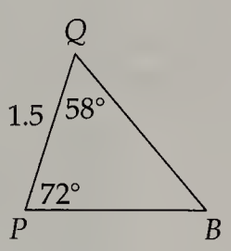
We'll tackle the cases in which $\triangle ABC$ is right or obtuse in Problem 4.17.

Problem 4.16: Let points P and Q be points on a shoreline that are 1.5 miles apart. Let B be the base of a radio tower on a distant island, and T be the top of the tower. A surveyor uses an angle measurement tool to determine that $\angle PQB = 58°$ and $\angle QPB = 72°$. She also measures that the angle of elevation from P to the top of the tower is 1°. How tall is the tower in feet? (There are 5280 feet in a mile.)

Solution for Problem 4.16: We first consider $\triangle BPT$ at right. We want BT, the height of the lighthouse, and the surveyor found that $\angle BPT = 1°$. If we could find any of the lengths of the sides of $\triangle BPT$, we could use trigonometry to find the remaining sides. Unfortunately, we don't yet know any of the sides of this triangle, so we turn to what we know about point Q.



We are given two angle measures and a side length of $\triangle PQB$, so we can find the third angle, and then use the Law of Sines to find the other sides. We have $\angle PBQ = 180° - 58° - 72° = 50°$. The side length we're most interested in is PB, since we can use that to find BT in the right triangle above. Applying the Law of Sines to $\triangle PQB$ gives

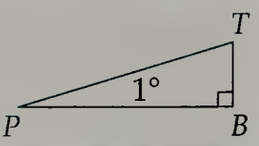


$$\frac{PB}{\sin \angle PQB} = \frac{PQ}{\sin \angle QBP},$$

so

$$PB = \frac{PQ \sin \angle PQB}{\sin \angle QBP} = \frac{1.5 \sin 58°}{\sin 50°} \approx 1.66.$$

Now we can find the height of the tower. In right triangle BPT, we have $\tan \angle BPT = \frac{BT}{PB}$. Since $PB \approx 1.66$ miles and $\angle BPT = 1°$, we have $BT = PB \tan \angle BPT \approx 0.0290$ miles. There are 5280 feet in one mile, so the height of the tower is $0.0290 \cdot 5280 \approx 153$ feet. □

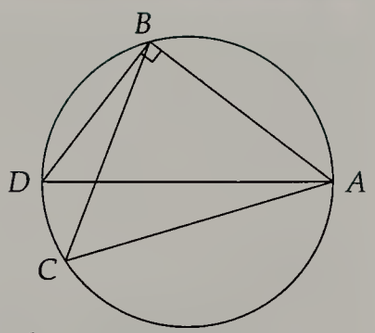


Problem 4.17: Prove the **Extended Law of Sines**, which states that if $a = BC$, $b = AC$, $c = AB$, and R is the radius of the circumcircle of $\triangle ABC$, then

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C} = 2R.$$

(Note: The circumcircle of a triangle is the circle that passes through all three vertices of the triangle.)

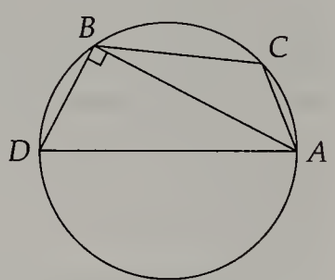
Solution for Problem 4.17: If $\triangle ABC$ is acute, then we already know that the first three ratios are equal, so we only have to prove that one of them equals $2R$. To get a clue where to start, we rewrite $\frac{c}{\sin C} = 2R$ as $\sin C = \frac{c}{2R}$. This suggests that we seek a right triangle with c as one leg length and $2R$ as the length of the hypotenuse. Therefore, we locate point D on the circle such that $\overline{AD}$ is a diameter (which has length $2R$), and we draw $\overline{AD}$ and $\overline{BD}$. Since $\angle ABD$ is inscribed in a semicircle, we have $\angle ABD = 90°$, so $\triangle ABD$ is a right triangle with leg $AB = c$ and hypotenuse $AD = 2R$. Then, we have $\sin \angle ADB = \frac{AB}{AD} = \frac{c}{2R}$, which means we only have to show that $\sin \angle ADB = \sin \angle C$ to finish. Since $\angle ADB$ and $\angle C$ are inscribed in the same arc of the circle, these two angles are equal, which means $\sin C = \sin \angle ADB = \frac{c}{2R}$, as desired.



Unfortunately, we're not quite finished. Our proof above assumes that $\angle C$ is acute. We have to tackle the cases of right angles and obtuse angles separately.

The proof is pretty straightforward for right angles; we can even use the diagram above. ADB is a right triangle with $\angle ABD = 90°$, and the circumradius of the triangle is $R = AD/2$, so $AD = 2R$. Since $\sin \angle ABD = \sin 90° = 1$, we also have $\frac{AD}{\sin \angle ABD} = AD$. Setting our two expressions for AD equal gives $\frac{AD}{\sin \angle ABD} = 2R$.

If $\angle C$ is obtuse, our first step is still to build right triangle ADB, as shown at right. As before, we have $\sin \angle ADB = \frac{AB}{AD} = \frac{c}{2R}$; however, we clearly don't have $\angle ADB = \angle C$ in this case. But we don't need the angles to be equal! We only need their sines to be equal to be able to conclude that $\sin C = \frac{c}{2R}$. Supplementary angles (angles that sum to 180°) have equal sines, so we wonder if $\angle C$ and $\angle ADB$ are supplementary. The measure of an angle inscribed in an arc of a circle is half the measure of the arc, so we have



$$\angle ADB + \angle C = \frac{\overset{\frown}{ACB}}{2} + \frac{\overset{\frown}{ADB}}{2} = \frac{\overset{\frown}{ACB} + \overset{\frown}{ADB}}{2}.$$

Putting the arcs $\overset{\frown}{ACB}$ and $\overset{\frown}{ADB}$ together gives us the whole circle, so the sum of their measures is 360°. Therefore, we have $\angle ADB + \angle C = 180°$, so $\angle ADB = 180° - \angle C$, and $\sin C = \sin(180° - \angle C) = \sin \angle ADB = \frac{c}{2R}$, as desired. □

Important: Let $a = BC$, $b = AC$, and $c = AB$ in $\triangle ABC$, and let R be the radius of the circumcircle of $\triangle ABC$ (that is, the radius of the circle that passes through all three vertices of $\triangle ABC$). The **Extended Law of Sines** tells us that

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C} = 2R.$$

Problem 4.18:

(a) Let $\triangle ABC$ and $\triangle DEF$ be triangles such that $AB = DE = 7$, $BC = EF = 5$, and $\angle CAB = \angle FDE = 30°$. Must $\sin \angle ACB = \sin \angle DFE$?

(b) Must $\triangle ABC$ and $\triangle DEF$ be congruent? If not, can we deduce anything about the relationship between $\angle ACB$ and $\angle DFE$?

Solution for Problem 4.18:

(a) Applying the Law of Sines to $\triangle ABC$ gives

$$\frac{AB}{\sin \angle ACB} = \frac{BC}{\sin \angle CAB} = \frac{5}{\sin 30°} = 10.$$

Therefore, we have

$$\sin \angle ACB = \frac{AB}{10} = \frac{7}{10}.$$

Similarly, applying the Law of Sines to $\triangle DEF$ gives

$$\frac{DE}{\sin \angle DFE} = \frac{EF}{\sin \angle FDE} = \frac{5}{\sin 30°} = 10,$$

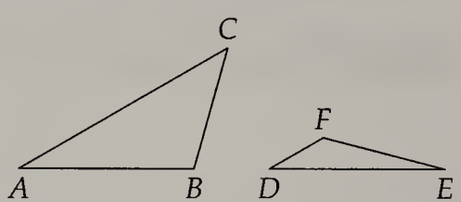
so

$$\sin \angle DFE = \frac{DE}{10} = \frac{7}{10}.$$

Therefore, we must have $\sin \angle ACB = \sin \angle DFE$.

(b) If $\triangle ABC \cong \triangle DEF$, then we must have $\angle ACB = \angle DFE$. At this point, all we know is that the sine of each of these angles equals $\frac{7}{10}$. However, there are two angles between 0° and 180° whose sine equals $\frac{7}{10}$, one acute angle and one obtuse angle. We can find these angles with a calculator. The acute angle is approximately 44.4°, and the obtuse angle is $180° - 44.4° = 135.6°$, since $\sin x = \sin(180° - x)$. We aren't given any information in the problem that prevents $\angle ACB$ or $\angle DFE$ from taking on either of these measures.

We might have $\angle ACB = 44.4°$ while $\angle DFE = 135.6$. Then, we have $\angle ABC = 180° - 44.4° - 30° = 105.6°$, while $\angle DEF = 180° - 135.6° - 30° = 14.4°$. In this case, the triangles are clearly not congruent. (The two triangles are drawn to scale at right.) All we can say is that if the triangles are not congruent, then $\angle ACB + \angle DFE = 180°$, since the sines of these two angles are equal but the angles are not equal.

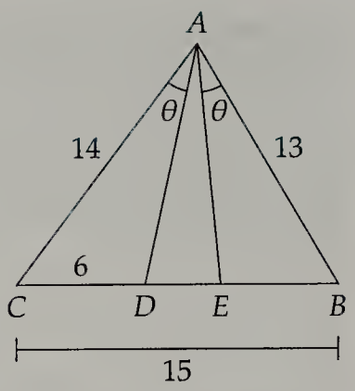


□

Sidenote: You probably learned various congruence theorems while studying geometry, such as Side-Angle-Side Congruence and Side-Side-Side Congruence. Problem 4.18 explains why there is no Side-Side-Angle Congruence Theorem. That is, if two sides and an angle other than the angle between these sides in one triangle are congruent to the corresponding sides and angle of another triangle, we *cannot* always conclude that the two triangles are congruent. (You'll see a special case in the Exercises in which Side-Side-Angle is enough to establish the congruence of two triangles, but usually it is not enough.)

Problem 4.19: In triangle ABC, $AB = 13$, $BC = 15$, and $CA = 14$. Point D is on $\overline{BC}$ with $CD = 6$. Point E is on $\overline{BC}$ such that $CE > CD$ and $\angle BAE = \angle CAD$. Find BE. *(Source: AIME)*

Solution for Problem 4.19: We let θ be the measure of equal angles $\angle BAE$ and $\angle CAD$. We need to use this angle equality, but we don't see any likely candidates for congruent triangles. We do have some equal side lengths, so perhaps the Law of Sines will allow us to use the angle equality somehow. Applying the Law of Sines to $\triangle ACD$ and $\triangle ABE$ gives:



$$\frac{BE}{\sin\theta} = \frac{AE}{\sin B} = \frac{AB}{\sin\angle AEB} \qquad \text{and} \qquad \frac{CD}{\sin\theta} = \frac{AD}{\sin C} = \frac{AC}{\sin\angle ADC}. \tag{4.1}$$

Dividing $\frac{BE}{\sin\theta} = \frac{AE}{\sin B}$ by $\frac{CD}{\sin\theta} = \frac{AD}{\sin C}$ cancels out $\sin\theta$ and leaves

$$\frac{BE}{CD} = \frac{AE}{AD}\cdot\frac{\sin C}{\sin B}.$$

We know CD, but we don't know any of the lengths on the right side. While we don't know the sines on the right, their ratio suggests applying the Law of Sines to $\triangle ABC$, since we know the side lengths of $\triangle ABC$. This gives us $\frac{AB}{\sin C} = \frac{AC}{\sin B}$, so $\frac{\sin C}{\sin B} = \frac{AB}{AC}$, and we know AB and AC. So, we now have

$$BE = \frac{AE}{AD}\cdot\frac{\sin C}{\sin B}\cdot CD = \frac{AE}{AD}\cdot\frac{AB}{AC}\cdot CD. \tag{4.2}$$

We know three of the lengths on the right side, but not AE or AD. Let's take a look at the third ratio in each of the original Law of Sines equations we found on line (4.1). Dividing $\frac{BE}{\sin\theta} = \frac{AB}{\sin\angle AEB}$ by $\frac{CD}{\sin\theta} = \frac{AC}{\sin\angle ADC}$ gives

$$\frac{BE}{CD} = \frac{AB}{AC}\cdot\frac{\sin\angle ADC}{\sin\angle AEB}.$$

Solving for BE gives us

$$BE = \frac{AB}{AC}\cdot\frac{\sin\angle ADC}{\sin\angle AEB}\cdot CD.$$

This looks familiar. The only difference between this and our final result on line (4.2) is that here we have $\frac{\sin\angle ADC}{\sin\angle AEB}$ in place of $\frac{AE}{AD}$. The Law of Sines together with the identity $\sin x = \sin(180^\circ - x)$ explains why these are equal. Applying the Law of Sines to $\triangle AED$ gives $\frac{AE}{\sin\angle ADE} = \frac{AD}{\sin\angle AED}$, so we have

$$\frac{AE}{AD} = \frac{\sin\angle ADE}{\sin\angle AED} = \frac{\sin(180^\circ - \angle ADC)}{\sin(180^\circ - \angle AEB)} = \frac{\sin\angle ADC}{\sin\angle AEB}.$$

This brings us right back to Equation (4.2).

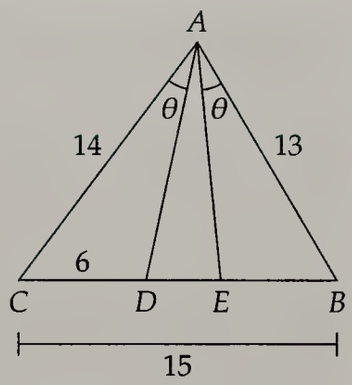
Concept: Most complex problems can be solved in more than one way. So, don't get trapped into thinking that you need to find *the* way to solve the problem.

We still have to find $\frac{AE}{AD}$ (or the corresponding ratio of sines). So far, we've made progress by focusing on each of the three little triangles, and on the large triangle $\triangle ABC$. The only triangles we haven't used yet are $\triangle ADB$ and $\triangle AEC$. Let's try the Law of Sines on these.

Concept: When stuck on a problem, experiment with parts of the problem you haven't used.

The Law of Sines applied to $\triangle AEC$ and $\triangle ADB$ gives us

$$\frac{AE}{\sin C} = \frac{EC}{\sin \angle EAC} = \frac{AC}{\sin \angle AEC} \quad \text{and} \quad \frac{AD}{\sin B} = \frac{BD}{\sin \angle BAD} = \frac{AB}{\sin \angle ADB}.$$



We start with the first two ratios in these equalities, because $\angle EAC = \angle BAD$ (since both equal $\theta + \angle DAE$). Dividing $\frac{AE}{\sin C} = \frac{EC}{\sin \angle EAC}$ by $\frac{AD}{\sin B} = \frac{BD}{\sin \angle BAD}$ gives

$$\frac{AE}{AD} \cdot \frac{\sin B}{\sin C} = \frac{EC}{BD} \cdot \frac{\sin \angle BAD}{\sin \angle EAC}.$$

We saw earlier that $\frac{\sin C}{\sin B} = \frac{AB}{AC}$, and $\angle BAD = \angle EAC$ means the sines on the right cancel. So, now we have

$$\frac{AE}{AD} = \frac{EC}{BD} \cdot \frac{\sin \angle BAD}{\sin \angle EAC} \cdot \frac{\sin C}{\sin B} = \frac{EC}{BD} \cdot \frac{AB}{AC}.$$

Success! We are given AB and AC, and we have $BD = BC - CD = 9$ and $EC = BC - BE = 15 - BE$. So, going back to our earlier expression for BE in Equation (4.2), we have

$$BE = \frac{AE}{AD} \cdot \frac{AB}{AC} \cdot CD = \frac{EC}{BD} \cdot \left(\frac{AB}{AC}\right)^2 \cdot CD = \frac{15 - BE}{9} \cdot \left(\frac{13}{14}\right)^2 \cdot 6.$$

Solving $BE = \frac{15-BE}{9} \cdot \left(\frac{13}{14}\right)^2 \cdot 6$ gives $BE = 2535/463$. □

Exercises

4.3.1 My house is on one side of a river and my neighbor's house is on the other. I'd like to figure out how far away his house is from mine, but I don't want to get wet, and we don't have a boat or a bridge. There's a well on my property that is 150 feet from my house. If I stand at my house and face the well, I must then turn 42° to my left to be facing my neighbor's house. If I stand at the well and face my house, I must turn 123° to my right to face my neighbor's house. To the nearest foot, what is the distance between my house and my neighbor's house?

4.3.2 The Leaning Tower of Alpine has been constructed 5° from perfectly vertical. When it casts a 20 foot shadow in the direction directly opposite the direction it leans, the angle of elevation from the tip of the shadow to the sun is 72°. How far is the top of the tower from the ground (to the nearest foot)?

4.3.3 Prove that if $AB > AC > BC$ in $\triangle ABC$, then $\angle C > \angle B > \angle A$. Is the converse true?

4.3.4 The perimeter of a triangle is numerically 12 times the average of the sines of the angles of the triangle. If one side of the triangle has length 2, what are the possible measures of the angle opposite that side?

4.3.5 In Problem 4.18, we encountered $\triangle ABC$ and $\triangle DEF$ such that $AB = DE = 7$, $BC = EF = 5$, and $\angle CAB = \angle FDE = 30°$, and showed that these two triangles need not be congruent. When I first created this problem to demonstrate why there is no "Side-Side-Angle Congruence Theorem," I had the 5 and 7 reversed, so that $AB = DE = 5$, $BC = EF = 7$, and $\angle CAB = \angle FDE = 30°$. Why did I have to change it?

4.3.6 What's wrong with this "proof" of the Angle-Angle Similarity Theorem:

Suppose we have $\triangle ABC$ and $\triangle DEF$ with $\angle A = \angle D$ and $\angle B = \angle E$. The Law of Sines gives us

$$\frac{BC}{\sin A} = \frac{AC}{\sin B} \quad \text{and} \quad \frac{EF}{\sin D} = \frac{DF}{\sin E}.$$

Therefore, we have $BC/AC = (\sin A)/(\sin B)$ and $EF/DF = (\sin D)/(\sin E)$. Since $\angle A = \angle D$ and $\angle B = \angle E$, we have $(\sin A)/(\sin B) = (\sin D)/(\sin E)$. So, we have

$$\frac{BC}{AC} = \frac{\sin A}{\sin B} = \frac{\sin D}{\sin E} = \frac{EF}{DF}.$$

Because $\angle A = \angle D$ and $\angle B = \angle E$, we have $\angle C = \angle F$. We can follow essentially the same steps as above to deduce that $BC : AC : AB = EF : DF : DE$.

4.3.7★ Distinct points A and B are on a semicircle with diameter $\overline{MN}$ and center C. The point P is on $\overline{CN}$ and $\angle CAP = \angle CBP = 10°$. If $\widehat{MA} = 40°$, then what is $\widehat{BN}$? *(Source: AHSME)*

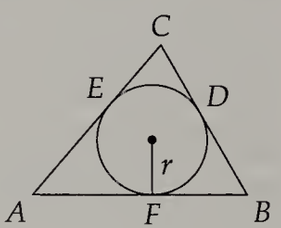
4.4 More Triangle Relationships

In addition to the Law of Sines and the Law of Cosines, there are many other triangle relationships that involve trigonometric functions of the angles of a triangle, or that can be proved using trig. In this section, we discover some of these relationships. More importantly, we learn how to develop them—there are far too many of these relationships to memorize them all. Instead, we learn how to derive them, so we can do so whenever we need them.

Before we get started, we'll review some notation we typically use with triangles, definitions of some of the key parts of a triangle, and a list of some of the basic relationships we will use.

We let the side lengths of $\triangle ABC$ be $a = BC$, $b = AC$, and $c = AB$. We denote the area of $\triangle ABC$ as $[ABC]$. The **semiperimeter** of a triangle is half the triangle's perimeter, and we usually use s to denote the semiperimeter. The circle that passes through all three vertices of $\triangle ABC$ is called the **circumcircle** of $\triangle ABC$ and its radius is called the **circumradius** of the triangle. The circle that is tangent to all three sides of a triangle is the **incircle** of the triangle; its radius is the **inradius** of the triangle. We usually use R for the circumradius of a triangle and r for the inradius of the triangle.

Two formulas that we proved in Art of Problem Solving's *Introduction to Geometry* that will be useful in this section are:



- Let F be the point where the incircle of $\triangle ABC$ is tangent to side $\overline{AB}$. Then, we have $AF = s - a$ and $BF = s - b$.
- $[ABC] = rs$.

You'll have a chance to prove these as Exercises.

Problems

Problem 4.20: Show that $[ABC] = \dfrac{abc}{4R}$. **Hints:** 194

Problem 4.21: Show that $b = c\cos A + a\cos C$ in acute triangle ABC. **Hints:** 280

Problem 4.22: Show that $\tan\dfrac{A}{2} = \dfrac{r}{s-a}$ and $\cos\dfrac{A}{2} = \sqrt{\dfrac{s(s-a)}{bc}}$. **Hints:** 176

Problem 4.23: Show that $\sin A + \sin B + \sin C = \frac{s}{R}$ in any triangle ABC.

Problem 4.24: Show that in any non-right triangle ABC, we have $\tan A + \tan B + \tan C = \tan A \tan B \tan C$. **Hints:** 203

Problem 4.25: In this problem, we prove **Stewart's Theorem**. Let point D be on side $\overline{BC}$, and let $d = AD$, $m = BD$, and $n = CD$. Stewart's Theorem states that

$$a(d^2 + mn) = mb^2 + nc^2.$$

(a) Which of the triangle relationships involving lengths is most likely to produce expressions like those in Stewart's Theorem?

(b) Apply your answer to part (a) to triangles ABD and ACD. Which angles of these two triangles will be most convenient to focus on, and why?

(c) Eliminate the trig functions from your results in part (b) and perform some clever algebra to prove that $a(d^2 + mn) = mb^2 + nc^2$.

Problem 4.20: Show that $[ABC] = \dfrac{abc}{4R}$.

Solution for Problem 4.20: Rather than start from scratch with a diagram, we try using formulas we already know.

Concept: Many triangle formulas can be proved by combining formulas you already know.

We look at the key features of $\frac{abc}{4R}$. First, we have a product of side lengths in the numerator. We already have an area formula with a product of side lengths:

$$[ABC] = \frac{1}{2}ab\sin C.$$

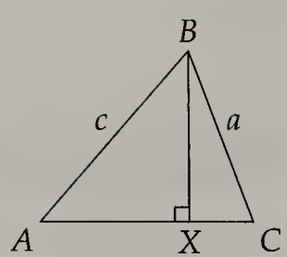
We also already have a formula involving R, the Extended Law of Sines, which tells us that $2R = \frac{c}{\sin C}$. Solving this for $\sin C$ gives $\sin C = \frac{c}{2R}$, and substituting this into $[ABC] = \frac{1}{2}ab\sin C$ gives $[ABC] = \frac{abc}{4R}$. □

Problem 4.21: Show that $b = c\cos A + a\cos C$ in acute triangle ABC.

Solution for Problem 4.21: We present two solutions:

Solution 1: Geometry. The two cosines suggest that we look for right triangles, one with $\angle A$ as one of the acute angles and one with $\angle C$. Drawing altitude $\overline{BX}$ accomplishes both. From right triangle BXA, we have $\cos A = \frac{AX}{AB} = \frac{AX}{c}$, so $AX = c\cos A$. Similarly, we have $CX = a\cos C$. Therefore, we have

$$b = AC = AX + CX = c\cos A + a\cos C.$$



Solution 2: Formula Bashing. We already have a formula that relates the cosine of an angle of a triangle to the side lengths of the triangle: the Law of Cosines. Solving for the cosine in the Law of Cosines gives

$$\cos A = \frac{b^2 + c^2 - a^2}{2bc}.$$

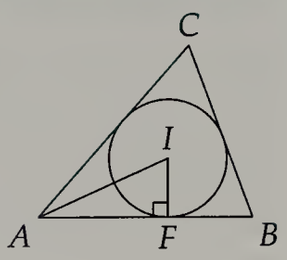
Doing the same for $\cos C$ and substituting into $c\cos A + a\cos C$ gives

$$c\cos A + a\cos C = c\cdot\frac{b^2 + c^2 - a^2}{2bc} + a\cdot\frac{a^2 + b^2 - c^2}{2ab} = \frac{b^2 + c^2 - a^2 + a^2 + b^2 - c^2}{2b} = b.$$

(Note that our second solution shows that this relationship holds for right and obtuse triangles, too.) □

Problem 4.22: Show that $\tan\frac{A}{2} = \frac{r}{s-a}$ and $\cos\frac{A}{2} = \sqrt{\frac{s(s-a)}{bc}}$.

Solution for Problem 4.22: We'll start with the formula for $\tan\frac{A}{2}$. Angle bisectors produce angles that are half the angles of the triangle, and the angle bisectors intersect at the center of the incircle. Because the incircle is tangent to all three sides of the triangle, we form a right triangle by connecting the incenter of the triangle to a vertex and to a neighboring point of tangency. The first desired formula jumps right out at us now. In right triangle IAF, we have $AF = s-a$, $IF = r$, and $\angle IAF = \frac{\angle A}{2}$, so $\tan\frac{A}{2} = \frac{r}{s-a}$.



Unfortunately, triangle IAF isn't going to give us $\cos\frac{A}{2}$ quickly, since we don't have a nice expression for IA. However, we can express $\cos\frac{A}{2}$ in terms of $\cos A$, and we can write $\cos A$ in terms of the side lengths of $\triangle ABC$ using the Law of Cosines. This plan might look daunting, but there are a couple reasons it also looks promising. First, the cosine half-angle formula introduces a square root, which we need. Second, the numerator of the desired expression has squared side lengths when we expand $s(s-a)$, and the Law of Cosines had squared side lengths. Applying the half-angle formula and then the Law of Cosines gives:

$$\begin{aligned}\cos\frac{A}{2} &= \sqrt{\frac{1+\cos A}{2}} = \sqrt{\frac{1+\frac{b^2+c^2-a^2}{2bc}}{2}} \\ &= \sqrt{\frac{2bc+b^2+c^2-a^2}{4bc}} = \sqrt{\frac{(b+c)^2-a^2}{4bc}} \\ &= \sqrt{\frac{(b+c+a)(b+c-a)}{4bc}} = \sqrt{\frac{(2s)(2s-2a)}{4bc}} = \sqrt{\frac{s(s-a)}{bc}}.\end{aligned}$$

□

Problem 4.23: Show that $\sin A + \sin B + \sin C = \frac{s}{R}$ in any triangle ABC.

Solution for Problem 4.23: Sines, side lengths, and the circumradius: this looks like a job for the Extended Law of Sines. Since

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C} = 2R,$$

we have

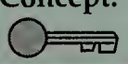
$$\sin A + \sin B + \sin C = \frac{a}{2R} + \frac{b}{2R} + \frac{c}{2R} = \frac{a+b+c}{2R} = \frac{s}{R}.$$

□

Problem 4.24: Show that in any non-right triangle ABC, we have $\tan A + \tan B + \tan C = \tan A \tan B \tan C$.

Solution for Problem 4.24: We start by asking ourselves why it's important that the three angles be angles of a non-right triangle.

Concept: Focusing on the key restrictions in a problem often leads to the solution.



In a right triangle, the tangent of one of the angles is undefined, so that explains why the triangle must be non-right. But what's important about A, B, and C being the angles of a triangle? The only thing we know from

the fact that the three angles are the angles of a triangle is $A + B + C = 180°$. This at least allows us to eliminate one of the variables in the identity. We have $C = 180° - (A + B)$, so $\tan C = \tan(180° - (A + B))$. Applying the identity $\tan(180° - x) = -\tan x$ gives $\tan C = \tan(180° - (A + B)) = -\tan(A + B)$, and now we have a clear path to the solution:

$$\tan C = -\tan(A + B) = -\frac{\tan A + \tan B}{1 - \tan A \tan B}. \tag{4.3}$$

Note that we cannot have $1 - \tan A \tan B = 0$ because this only occurs if $\tan A = \frac{1}{\tan B} = \cot B$. This only holds in a triangle if $A + B = 90°$. However, we are told that the triangle is not a right triangle, so we cannot have $A + B = 90°$.

Multiplying both ends of Equation (4.3) by $1 - \tan A \tan B$ and rearranging gives the desired

$$\tan A + \tan B + \tan C = \tan A \tan B \tan C.$$

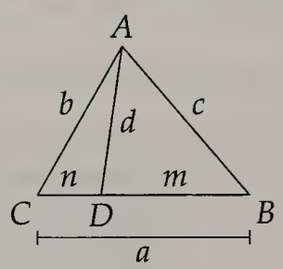
□

Problem 4.25: Let point D be on side $\overline{BC}$, and let $d = AD$, $m = BD$, and $n = CD$. **Stewart's Theorem** states that

$$a(d^2 + mn) = mb^2 + nc^2.$$

Prove Stewart's Theorem.

Solution for Problem 4.25: The desired equation has multiple squares of side lengths, just like the Law of Cosines. We'll likely have to apply the Law of Cosines to both $\triangle ABD$ and $\triangle ACD$, since we need equations involving m, n, and d. We focus on $\angle ADC$ and $\angle ADB$, because these angles are supplementary, which allows us to easily relate their cosines: $\cos \angle ADC = \cos(180° - \angle ADB) = -\cos \angle ADB$. From the Law of Cosines, we have



$$\cos \angle ADC = \frac{d^2 + n^2 - b^2}{2dn},$$
$$\cos \angle ADB = \frac{d^2 + m^2 - c^2}{2dm}.$$

Therefore, $\cos \angle ADC = -\cos \angle ADB$ gives us

$$\frac{d^2 + n^2 - b^2}{2dn} = -\frac{d^2 + m^2 - c^2}{2dm}.$$

Multiplying both sides by $2dmn$ gives

$$d^2m + mn^2 - b^2m = -(d^2n + m^2n - c^2n).$$

Rearranging this equation gives $d^2m + d^2n + mn^2 + m^2n = b^2m + c^2n$. Factoring gives $d^2(m+n) + mn(m+n) = b^2m + c^2n$, and because $m + n = a$, we have the desired $a(d^2 + mn) = mb^2 + nc^2$. □

Sidenote: For some people, Stewart's Theorem is easier to remember as

$$dad + man = bmb + cnc.$$

Most of the relationships we discovered in this section are not that important, which is why they are not in Important boxes. However, the two key strategies we used to find them are important:

Concepts: Two strategies for proving triangle identities are:

- Use trigonometric and geometric identities that you have already proved. Use parts of the desired identity as clues for which known identities you should try to use.
- Use geometric tools in combination with trigonometry, often by building right triangles.

Exercises

4.4.1 Let F be the point where the incircle of $\triangle ABC$ is tangent to side $\overline{AB}$. Show that $AF = s - a$ and $BF = s - b$.

4.4.2 Show that $[ABC] = rs$.

4.4.3 Complete our geometric proof that $b = c\cos A + a\cos C$ by addressing the cases in which $\triangle ABC$ is right or obtuse.

4.4.4 Show that in $\triangle ABC$, we have $\sin A \sin B \sin C = \dfrac{[ABC]}{2R^2}$.

4.4.5

(a) Use the identity we found for $\cos\dfrac{A}{2}$ to show that $\sin\dfrac{A}{2} = \sqrt{\dfrac{(s-b)(s-c)}{bc}}$.

(b) Prove that $[ABC] = \sqrt{s(s-a)(s-b)(s-c)}$. (This is known as **Heron's formula**. We present a dazzling proof of Heron's formula with complex numbers on page 484.)

(c) Show that $R = \dfrac{abc}{\sqrt{(a+b+c)(a-b+c)(a+b-c)(b+c-a)}}$.

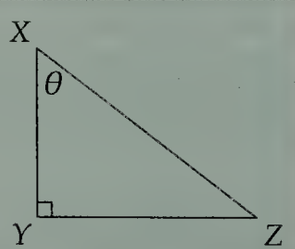
4.4.6 Let M be the midpoint of $\overline{BC}$. Find the length of median $\overline{AM}$ in terms of a, b, and c.

4.4.7⋆ Let $\triangle ABC$ be an acute triangle whose altitudes meet at point H. Show that $AH = 2R\cos A$. **Hints:** 279

4.4.8⋆ Show that in $\triangle ABC$, if $a = BC$ and $b = AC$, then we have $\dfrac{\tan\frac{A-B}{2}}{\tan\frac{A+B}{2}} = \dfrac{a-b}{a+b}$. This is sometimes called the **Law of Tangents**. **Hints:** 30

4.5 Summary

Important: If triangle XYZ is a right triangle with $\angle XYZ = 90°$ and $\angle X = \theta$, then



$$\sin\theta = \frac{YZ}{XZ}, \qquad \cos\theta = \frac{YX}{XZ}, \qquad \tan\theta = \frac{YZ}{XY}.$$

Another way of viewing this is: If θ is the measure of an acute angle in a right triangle, then we have

$$\sin\theta = \frac{\text{opposite leg}}{\text{hypotenuse}}, \qquad \cos\theta = \frac{\text{adjacent leg}}{\text{hypotenuse}}, \qquad \tan\theta = \frac{\text{opposite leg}}{\text{adjacent leg}}.$$

Important: If $a = BC$, $b = AC$, and $[ABC]$ is the area of $\triangle ABC$, then

$$[ABC] = \frac{1}{2}ab\sin C.$$

Important: Let $a = BC$, $b = AC$, and $c = AB$ in $\triangle ABC$. The **Law of Cosines** states that

$$c^2 = a^2 + b^2 - 2ab\cos C.$$

Important: Let $a = BC$, $b = AC$, and $c = AB$ in $\triangle ABC$, and let R be the radius of the circumcircle of $\triangle ABC$ (the radius of the circle that passes through all three vertices of $\triangle ABC$). The **Extended Law of Sines** tells us that

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C} = 2R.$$

Things To Watch Out For!

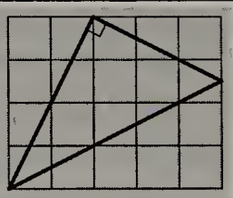
When proving a statement for all possible triangles, make sure your proof is valid for all possible triangles. Specifically, you may have to address right and obtuse triangles as separate cases from acute triangles.

Problem Solving Strategies

Concepts:

- Many geometry problems can be solved by assigning variables to lengths or angles in the problems, and then finding other lengths or angles in terms of the variables. The goal is usually to find a way to build an equation that you can solve.
- Area can be a surprisingly powerful problem solving tool, even in problems that don't appear to be about area.
- Once you think you've found a new formula, check your work by trying the formula on specific examples you have solved without the formula.
- Solving a specific example of a general problem can often provide a guide for solving the general problem.
- When stuck on a problem, focus on information you haven't used yet.
- If a tactic gives some information in a problem, but doesn't completely solve it, then try applying that tactic to the problem again in another way.
- Most complex problems can be solved in more than one way. So, don't get trapped into thinking that you need to find *the* way to solve the problem.
- Focusing on the key restrictions in a problem often leads to the solution.

Extra! On page 135, we asked you to find a 3-4-5 right triangle in a problem. The 3-4-5 triangle is in bold at right. The lower left square in the grid corresponds to the square in the original problem.



REVIEW PROBLEMS

4.26 In right triangle STU, we have $\angle S = 3\angle T$ and $\angle U = 90°$. Find ST/TU to the nearest hundredth.

4.27 Find all possible values of $\angle PQR$ if $PQ = 12$, $QR = 4\sqrt{3}$, and the area of $\triangle PQR$ is $12\sqrt{6}$.

4.28 If $JK = 12$, $KL = 7$, and $\angle JLK = 90°$, then what is $\angle JKL$ to the nearest degree?

4.29 Let the graphs of $2x - 3y = 7$ and $3x - y = 9$ be k and ℓ, respectively.

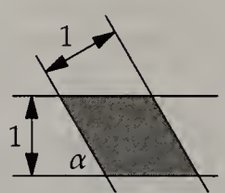
(a) To the nearest degree, find the acute angle between k and the x-axis.

(b) To the nearest degree, find the acute angle between ℓ and the x-axis.

(c) To the nearest degree, find the acute angle between k and ℓ.

(d)★ Find an expression for the tangent of an angle between two lines with slopes m_1 and m_2, where $m_1 \neq m_2$. Test your result with k and ℓ.

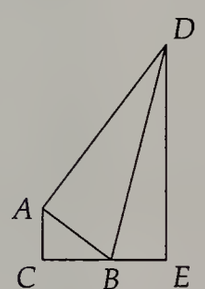
4.30 Two strips of width 1 overlap at an angle of α as shown. What is the area of the overlap in terms of α? *(Source: AHSME)*



4.31 A boat sails 9 miles per hour along the latitude 30° N. Assuming that the Earth is a sphere with radius 3950 miles, how long will it take the boat to sail from longitude 19° W to 46° W (to the nearest hour)?

4.32 M is the midpoint of side $\overline{AB}$ of $\triangle ABC$. If $\angle ABC = 90°$, and $\sin \angle MCB = 0.8$, then what is $\cos \angle ACB$?

4.33 Triangle ABC has a right angle at C, $AC = 3$ and $BC = 4$. Triangle ABD has a right angle at A and $AD = 12$. Points C and D are on opposite sides of $\overline{AB}$. The line through D parallel to $\overline{AC}$ meets $\overline{CB}$ extended at E. Find DE/DB. *(Source: AHSME)*



4.34 In $\triangle GHI$, we have $GH = HI = 9$ and $GI = 6$.

(a) Find $\tan \angle GIH$.

(b) Find $\tan \angle GHI$.

(c) Find $\cos \angle GHI$.

4.35 When I look due west, I look straight at a mountain peak that I know is 4.6 miles away. If I turn northward 37°, I will be looking directly at a mountain peak that is 5.8 miles away. To the nearest tenth of a mile, how far apart are the mountain peaks?

4.36 What's wrong with the following proof of the Pythagorean Theorem:

> In the text, we proved that if $a = BC$, $b = AC$, and $c = AB$, then $c^2 = a^2 + b^2 - 2ab\cos C$. If $\angle C = 90°$, then we have $\cos C = 0$. Therefore, if $\angle C = 90°$, we have $c^2 = a^2 + b^2$.

4.37 Find the angles of a triangle with side lengths 2, 2, and $\sqrt{6} - \sqrt{2}$.

4.38 In $\triangle ABC$, $AB = 5$, $BC = 7$, $AC = 9$, and D is on $\overline{AC}$ with $BD = 5$. Find AD/DC. *(Source: AHSME)*

4.39 Is it possible for the three side lengths of a triangle to form an arithmetic sequence with common difference 1, and have one angle with measure 60°? Why or why not?

4.40 We define the **dihedral angles** between two intersecting planes as follows. Suppose line k is the intersection of the planes, and point A is on k. Let point B be in one plane and C in the other such that both $\overline{AB}$ and $\overline{AC}$ are perpendicular to k. Then, $\angle BAC$ and its supplement are the dihedral angles between the planes. (There are two dihedral angles between two intersecting planes, just as there are two angles between two intersecting lines.)

A regular tetrahedron is a triangular pyramid in which all four faces are equilateral triangles. Find the acute dihedral angle between two faces of a regular tetrahedron to the nearest tenth of a degree.

4.41 In $\triangle PQR$, we have $PQ = 4$, $QR = 6$, and $\angle P = 2\angle R$. Find $\cos \angle R$.

4.42 Does there exist a triangle ABC such that $AB = 3$, $BC = 8$, $\angle A = 73°$, and $\angle C = 57°$? Why or why not?

4.43 A plane takes off, flies in straight line for 350 miles, then turns 20° and flies straight for another 2 hours before landing at an airport that is 800 miles from where the plane took off. How long did the whole flight take if the plane flew at a constant speed throughout the flight?

4.44 The base of an isosceles triangle is 6 inches and one of the equal sides is 12 inches. What is the radius of the circle through the vertices of the triangle? *(Source: AHSME)*

4.45 Show that in $\triangle ABC$, if $AB \cos B = AC \cos C$, then $\angle B = \angle C$.

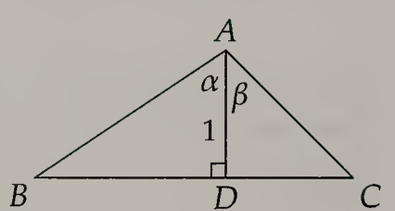
4.46 We showed in the text that there is no such thing as the "Side-Side-Angle Congruence Theorem." However, suppose $AB = DE = 6$, $AC = DF = 12$, and $\angle BCA = \angle EFD = 30°$. Use the Law of Sines to explain why we can conclude that $\triangle ABC \cong \triangle DEF$ in this special case. What makes this case special?

4.47 The **Angle Bisector Theorem** states that if D is on side $\overline{BC}$ of $\triangle ABC$ such that $\angle BAD = \angle CAD$, then $AB/BD = AC/CD$. Prove the Angle Bisector Theorem.

Challenge Problems

4.48 In the diagram at right, $\overline{AD}$ is an altitude of $\triangle ABC$ with length 1, and angles B and C are acute. Use the diagram at right to prove the angle sum identity for sine,

$$\sin(\alpha + \beta) = \sin\alpha \cos\beta + \sin\beta \cos\alpha.$$



Hints: 278

4.49 Show that in $\triangle ABC$, we have $4\cos\frac{A}{2}\cos\frac{B}{2}\cos\frac{C}{2} = \frac{s}{R}$.

4.50 Let $\triangle ABC$ be an acute triangle whose altitudes meet at point H, and let X be on $\overline{BC}$ such that $\overline{AX} \perp \overline{BC}$. Show that $HX = 2R \cos B \cos C$. **Hints:** 63

4.51 In triangle ABC, D is on $\overline{AC}$ and F is on $\overline{BC}$. Also, $\overline{AB} \perp \overline{AC}$, $\overline{AF} \perp \overline{BC}$, and $BD = DC = FC = 1$. Find AC. *(Source: AHSME)* **Hints:** 49

4.52 In quadrilateral $ABCD$, $BC = 8$, $CD = 12$, $AD = 10$, and $\angle A = \angle B = 60°$. Find AB. *(Source: AIME)* **Hints:** 129, 101

4.53 A right circular cone has a base with radius 600 and height $200\sqrt{7}$. An ant starts at a point on the surface of the cone whose distance from the vertex of the cone is 125, and crawls along the surface of the cone to a point on the exact opposite side of the cone whose distance from the vertex is $375\sqrt{2}$. Find the least distance that the ant could have crawled. *(Source: AIME)* **Hints:** 259

4.54 In $\triangle ABC$, the angles A, B, and C satisfy the equation $\cos A \cos B + \sin A \sin B \sin C = 1$. Determine all possible values of $\angle C$. *(Source: CEMC)* **Hints:** 65

4.55 Consider a pyramid $PABCD$ whose base $ABCD$ is square and whose vertex P is equidistant from A, B, C, and D. Show that if $AB = 1$ and $\angle APB = 2\theta$, then the volume of the pyramid is $\sqrt{\cos 2\theta}/(6 \sin \theta)$. *(Source: AHSME)*

4.56 Points A, B, and C on a circle of radius r are situated so that $AB = AC$, $AB > r$, and the length of minor arc $\widehat{BC}$ is r. Show that $AB/BC = \frac{1}{2} \csc \frac{1}{4}$. *(Source: AHSME)* **Hints:** 266

4.57 The lengths of the sides of a triangle are consecutive integers, and the largest angle is twice the smallest angle. What is the cosine of the smallest angle? *(Source: AHSME)* **Hints:** 248

4.58 Point D is on side $\overline{CB}$ of triangle ABC. If $\angle CAD = \angle DAB = 60°$, $AC = 3$, and $AB = 6$, then what is the length of $\overline{AD}$? *(Source: AHSME)*

4.59 An observer walks along a level road directly towards an elevated object P. He takes observations at three points A, B, and C, in order, and notes that the angles of elevation of P are θ, 2θ, and 3θ, respectively. Prove that $AB > 2BC$. *(Source: CEMC)* **Hints:** 238, 133

4.60 Point D is on $\overline{BC}$ such that $\angle BAD = \angle CAD$. Find AD in terms of a, b, and c, where $a = BC$, $b = AC$, and $c = AB$.

4.61 Triangle ABC is isosceles with $AB = AC$ and altitude $AM = 11$. Suppose that there is a point D on $\overline{AM}$ with $AD = 10$ and $\angle BDC = 3\angle BAC$. Find the perimeter of $\triangle ABC$. *(Source: AIME)* **Hints:** 110, 92

4.62★ Find $\sin 18°$. (See if you can find both a geometric and an algebraic solution.) **Hints:** 157, 90

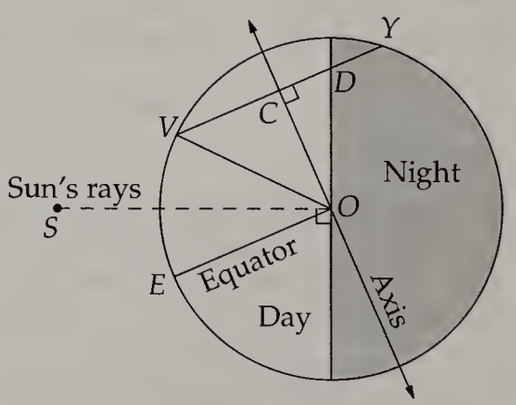
4.63 Let a, b, c be the three sides of a triangle, and let α, β, γ, respectively, be the angles opposite them. If $a^2 + b^2 = 1989c^2$, find

$$\frac{\cot \gamma}{\cot \alpha + \cot \beta}.$$

(Source: AIME) **Hints:** 159

4.64 Show that if A, B, and C are the angles of a triangle, then $\tan \frac{A}{2} \tan \frac{B}{2} + \tan \frac{B}{2} \tan \frac{C}{2} + \tan \frac{C}{2} \tan \frac{A}{2} = 1$.

4.65★ The axis of the Earth makes an angle of 66.5° with the plane in which it orbits the sun. The latitude of Vancouver is 49° N, which means that if we form an angle by connecting Vancouver to the center of the Earth, and then connecting the center of the Earth to a point on the equator with the same longitude as Vancouver, the angle has measure 49°. (This is $\angle VOE$ at right.)



At right is a cross-section of the Earth that includes Vancouver and the Earth's axis. Vancouver is point V, point E is on the equator, point O is the center of the Earth, S is the Sun, and $\overline{VY}$ is perpendicular to the Earth's axis. The summer solstice is the day of the year with the most hours of sunlight in the Northern hemisphere. Find the number of hours of sunlight Vancouver receives on the summer solstice. **Hints:** 219, 27, 171

4.66 Triangle ABC is a right triangle with $AC = 7$, $BC = 24$, and right angle at C. Point M is the midpoint of AB, and D is on the same side of line AB as C so that $AD = BD = 15$. Find the area of $\triangle CDM$. *(Source: AIME)*

4.67★ A circle is inscribed in quadrilateral $ABCD$, tangent to $\overline{AB}$ at P and to $\overline{CD}$ at Q. Given that $AP = 19$, $PB = 26$, $CQ = 37$, and $QD = 23$, find the square of the radius of the circle. *(Source: AIME)* **Hints:** 78

4.68★ Show that if A, B, and C are the angles of a triangle, then

$$\cos^2 A + \cos^2 B + \cos^2 C + 2\cos A \cos B \cos C = 1.$$

Hints: 68, 21

4.69★ .

(a) Prove that if $ABCD$ is a **cyclic quadrilateral** (a quadrilateral that can be inscribed in a circle), then the area of $ABCD$ is $\sqrt{(s-a)(s-b)(s-c)(s-d)}$, where a, b, c, and d are the side lengths of $ABCD$ and $s = (a+b+c+d)/2$. (This result is known as **Brahmagupta's formula**.) **Hints:** 260, 3

(b)★ Prove that the area of any quadrilateral $ABCD$ is

$$\sqrt{(s-a)(s-b)(s-c)(s-d) - abcd\cos^2\left(\frac{A+C}{2}\right)}.$$

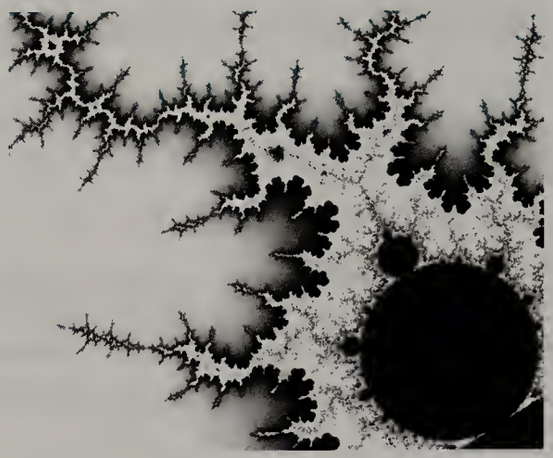
(This result is known as **Bretschneider's formula**.) **Hints:** 186, 169, 286

4.70★ In parallelogram $ABCD$, let O be the intersection of diagonals $\overline{AC}$ and $\overline{BD}$. Angles CAB and DBC are each twice as large as angle DBA, and angle ACB is r times as large as angle AOB. Find r. *(Source: AIME)* **Hints:** 7

4.71★ Show that in $\triangle ABC$, we have

$$BC^3\cos(B-C) + CA^3\cos(C-A) + AB^3\cos(A-B) = 3(BC)(CA)(AB).$$

Hints: 265



Mathematics is the supreme judge; from its decisions there is no appeal. – Tobias Dantzig

CHAPTER 5

Parameterization and Trigonometric Coordinate Systems

5.1 Parameterization

Imagine we wished to track the path of a particle traveling within the Cartesian plane. Suppose that we want to know more than just the path it takes, but we also want to know where it is at each moment in time. Then, saying that the particle travels along the graph of an equation like $y = x + 2$ won't do—this doesn't tell us where the particle is at each moment in time. Instead, we can use equations called **parametric equations**.

Rather than having a single equation that relates x and y, we can define a relationship between x and y using an intermediary variable called a **parameter**. We do so by expressing x and y as functions of the parameter, meaning we relate x and y with equations of the form $x = f(t)$, $y = g(t)$, where t is the parameter. (We often use t as the variable for a parameter because we frequently use a parameter to represent time.) We refer to the equations $x = f(t)$, $y = g(t)$ as parametric equations. When we graph parametric equations $x = f(t)$, $y = g(t)$, we graph all points (x, y) such that there is a value of t for which $x = f(t)$ and $y = g(t)$. If the possible values of the parameter are not specified, then it is implied that the parameter can take on any real value for which the equations are defined.

For example, we can use the parametric equations

$$x = 2 - 3t,$$
$$y = 1 + 2t,$$

to represent the location of a particle that is at $(x, y) = (2, 1)$ at time $t = 0$ and moves 3 to the left and up 2 for each unit of time. When $t = 4$ in these parametric equations, we have $x = -10$ and $y = 9$, so $(x, y) = (-10, 9)$ is on the graph of the parametric equations. If we let $t = -1$, we have $(x, y) = (5, -1)$, so if we leave t unrestricted, then $(x, y) = (5, -1)$ is on the graph. To restrict the graph to only those points the particle will pass through from the time it starts at $(2, 1)$, we would restrict the parameter to $t \geq 0$.

Problems

Problem 5.1: In this problem, we find the graph of the parametric equations $x = 2 - 3t$, $y = 1 + 2t$.

(a) Choose several values of t and compute (x, y) for each value. Plot the resulting points. What does it look like the graph of all such points will be?

(b) Eliminate t from the parametric equations $x = 2 - 3t$, $y = 1 + 2t$ to find an equation in which the only variables are x and y. Show that for every pair (x_1, y_1) that satisfies this new equation, we can find a value of t in the parametric equations that gives $(x, y) = (x_1, y_1)$.

(c) Suppose we limit t to the interval $[4, 8]$. How does this affect the graph?

Problem 5.2: Find the graph of the parametric equations $x = 1 - t$, $y = 2t^2 - t$.

Problem 5.3: Find the graph of the parametric equations $x = 3 + 2\cos t$, $y = 4 + 2\sin t$ for $0 \leq t < 2\pi$.

Problem 5.4: Find the graph of the parametric equations $x = \sin\theta$, $y = 3 - 2\cos 2\theta$. (Be careful!)

Problem 5.5:

(a) Find parametric equations for x and y such that the graph of the parametric equations is the same as the graph of $x^2 + y^2 = 1$.

(b) Is your answer to part (a) unique? Are there any other parametric equations that produce the same graph?

(c) Suppose a particle starts at the point $(0, -2)$, and moves at a constant speed clockwise around the origin along the graph of $x^2 + y^2 = 4$. If the particle completes a full revolution in 4 seconds, then find parametric equations for the location of the particle in terms of t, where t is the time in seconds since the particle started moving.

Problem 5.6: Find parametric equations for x and y such that $x^2 - 3y^2 = 1$.

Problem 5.7: I'm riding my bike at a constant speed of 20 miles per hour on flat land when I ride over a dot of fresh yellow paint, thereby getting a yellow spot on my front wheel, which is 2 feet in diameter. I continue riding my bike in a straight line at 20 miles per hour. When t seconds have passed since running over the dot, how many feet is the yellow spot on my tire from the yellow dot I ran over?

Problem 5.1: In this problem, we find the graph of the parametric equations $x = 2 - 3t$, $y = 1 + 2t$.

(a) Choose several values of t and compute (x, y) for each value. Plot the resulting points. What does it look like the graph of all such points will be?

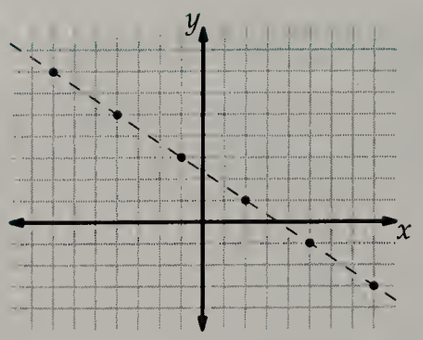
(b) Eliminate t from the parametric equations $x = 2 - 3t$, $y = 1 + 2t$ to find an equation in which the only variables are x and y. Show that for every pair (x_1, y_1) that satisfies this new equation, we can find a value of t in the parametric equations that gives $(x, y) = (x_1, y_1)$.

(c) Suppose we limit t to the interval $[4, 8]$. How does this affect the graph?

Solution for Problem 5.1:

(a) At left atop the following page we have a table of values of t and corresponding values of x and y. At right, we have graphed several of the resulting points (x, y) on the Cartesian plane.

t	$x = 2 - 3t$	$y = 1 + 2t$
-4	14	-7
-3	11	-5
-2	8	-3
-1	5	-1
0	2	1
1	-1	3
2	-4	5
3	-7	7



The points on the graph look collinear. It looks like the graph of the parametric equations is a line.

(b) We can confirm our guess from part (a) by eliminating the parameter from the parametric equations. Adding 2 times $x = 2 - 3t$ to 3 times $y = 1 + 2t$ gives $2x + 3y = 7$. The graph of this equation is indeed a line.

Concept: We can sometimes better understand a set of parametric equations by eliminating the parameter from the equations.

Unfortunately, this only tells us that every point in the graph of the parametric equations is in the graph of $2x + 3y = 7$. In order to conclude that the graphs are the same, we must show that every point in the graph of $2x + 3y = 7$ is in the graph of the parametric equation. Fortunately, that's not too hard. Suppose that (x_1, y_1) is in the graph of $2x + 3y = 7$, so we have $2x_1 + 3y_1 = 7$. To show that (x_1, y_1) is in the graph of the parametric equations, we must show that there is a value of t that gives $(x, y) = (x_1, y_1)$. To find that t, we solve the parametric equation $x = 2 - 3t$ for t, which gives $t = (2 - x)/3$. So, we let $t = (2 - x_1)/3$ in the parametric equations, and we have

$$x = 2 - 3 \cdot \frac{2 - x_1}{3} = x_1,$$
$$y = 1 + 2 \cdot \frac{2 - x_1}{3} = \frac{7}{3} - \frac{2}{3}x_1 = \frac{7}{3} - \frac{2}{3} \cdot \frac{7 - 3y_1}{2} = y_1,$$

which means (x_1, y_1) is in the graph of the parametric equations. Therefore, every point in the graph of $2x + 3y = 7$ is in the graph of the parametric equations. We conclude that the graph of $x = 2 - 3t$, $y = 1 + 2t$ is the same as the graph of $2x + 3y = 7$.

WARNING!! If parametric equations $x = f(t)$, $y = g(t)$ satisfy an equation in terms of x and y, we only know that the graph of the parametric equations is *part of* the graph of the non-parametric equation. To show that the graph of the parametric equations is the same as the graph of the non-parametric equation, we must show that every point in the graph of the non-parametric equation is in the graph of the parametric equations. We can usually do so by showing that for each point (x_1, y_1) on the graph of the non-parametric equation, there is a value of t such that $x_1 = f(t)$, $y_1 = g(t)$.

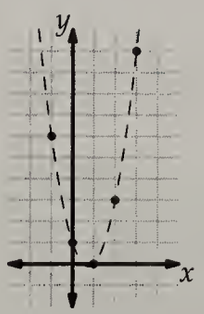
(c) When $t = 4$, we have $(x, y) = (-10, 9)$, and when $t = 8$, we have $(x, y) = (-22, 17)$, so the graph of the parametric equations when t is restricted to the interval $[4, 8]$ is the line segment from $(-10, 9)$ to $(-22, 17)$. To see that every point on this segment is on the graph, and no other points are, we can start from $4 \le t \le 8$, multiply by -3 to produce $-12 \ge -3t \ge -24$, and then add 2 to get $-10 \ge 2 - 3t \ge -22$. Since $x = 2 - 3t$, we have $-10 \ge x \ge -22$. So, every point on the line from part (b) for which $-10 \ge x \ge -22$ is produced by values of t in the interval $[4, 8]$, and no other values of x are possible.

□

Problem 5.2: Find the graph of the parametric equations $x = 1 - t$, $y = 2t^2 - t$.

Solution for Problem 5.2: We start as in Problem 5.1, choosing values of t and solving for x and y.

t	$x = 1 - t$	$y = 2t^2 - t$
-3	4	21
-2	3	10
-1	2	3
0	1	0
1	0	1
2	-1	6
3	-2	15



The t^2 in the parametric equation for y suggests the graph might be a parabola, and the dashed parabola in the diagram seems to fit the data nicely. To see that the graph is indeed a parabola, we try to relate x directly to y. We can't use elimination, as we did in Problem 5.1, but substitution does the job. From $x = 1 - t$, we have $t = 1 - x$, so

$$y = 2t^2 - t = 2(1 - x)^2 - (1 - x) = 2x^2 - 4x + 2 - 1 + x = 2x^2 - 3x + 1.$$

As in the previous problem, we still must show that every point in the graph of $y = 2x^2 - 3x + 1$ is in the graph of the parametric equations. For any point (x_1, y_1) in the graph of $y = 2x^2 - 3x + 1$, we have $y_1 = 2x_1^2 - 3x_1 + 1$. Solving the parametric equation $x = 1 - t$ for t gives $t = 1 - x$. So, we let $t = 1 - x_1$ in the parametric equations, and we have

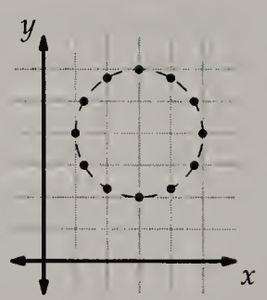
$$x = 1 - t = 1 - (1 - x_1) = x_1,$$
$$y = 2t^2 - t = 2(1 - x_1)^2 - (1 - x_1) = 2x_1^2 - 3x_1 + 1 = y_1,$$

which means that (x_1, y_1) is in the graph of the parametric equations, as well. Therefore, the graph of the parametric equations is the same as the graph of $y = 2x^2 - 3x + 1$, which is the dashed parabola shown above. □

Problem 5.3: Find the graph of the parametric equations $x = 3 + 2\cos t$, $y = 4 + 2\sin t$ for $0 \le t < 2\pi$.

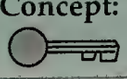
Solution for Problem 5.3: We choose values of t for which we can evaluate sine and cosine easily.

t	$x = 3 + 2\cos t$	$y = 4 + 2\sin t$	t	$x = 3 + 2\cos t$	$y = 4 + 2\sin t$
0	5	4	π	1	4
$\pi/6$	$3 + \sqrt{3}$	5	$7\pi/6$	$3 - \sqrt{3}$	3
$\pi/3$	4	$4 + \sqrt{3}$	$4\pi/3$	2	$4 - \sqrt{3}$
$\pi/2$	3	6	$3\pi/2$	3	2
$2\pi/3$	2	$4 + \sqrt{3}$	$5\pi/3$	4	$4 - \sqrt{3}$
$5\pi/6$	$3 - \sqrt{3}$	5	$11\pi/6$	$3 + \sqrt{3}$	3



It looks like the graph is the circle shown at right. We can't use either the elimination tactic of Problem 5.1 or the straightforward substitution tactic of Problem 5.2 to turn the parametric equations into a single equation with only x and y. So, we look for other strategies for eliminating the $\cos t$ and $\sin t$ terms, which brings us to the trigonometric identity $\sin^2 t + \cos^2 t = 1$.

Concept: Trig identities can be very useful in working with parametric equations.



Solving $x = 3 + 2\cos t$ and $y = 4 + 2\sin t$ for $\cos t$ and $\sin t$ give $\cos t = (x - 3)/2$ and $\sin t = (y - 4)/2$. Substituting

these into $\sin^2 t + \cos^2 t = 1$ gives

$$\left(\frac{x-3}{2}\right)^2 + \left(\frac{y-4}{2}\right)^2 = 1.$$

Multiplying both sides by 2^2 gives

$$(x-3)^2 + (y-4)^2 = 2^2.$$

The graph of this equation is a circle with center (3, 4) and radius 2, as expected, and as depicted above.

We're not finished yet! We have only shown that every point in the graph of the parametric equations is on the circle. We still have to show that every point on the circle is in the graph of the parametric equations.

We know one circle that is related to cosine and sine—we use the unit circle to define cosine and sine. From this definition, the unit circle is the graph of the parametric equations $x = \cos t$, $y = \sin t$. Multiplying the right sides of these equations simply scales every point away from the origin by a factor of 2, so the graph of $x = 2\cos t$, $y = 2\sin t$ is a circle centered at the origin with radius 2. We produce the graph of $x = 3 + 2\cos t$, $y = 4 + 2\sin t$ by translating each point on the graph of $x = 2\cos t$, $y = 2\sin t$ to the right 3 and up 4. Therefore, the graph of $x = 3 + 2\cos t$, $y = 4 + 2\sin t$ is the full circle centered at (3, 4) with radius 2. □

Problem 5.4: Find the graph of the parametric equations $x = \sin\theta$, $y = 3 - 2\cos 2\theta$. (Be careful!)

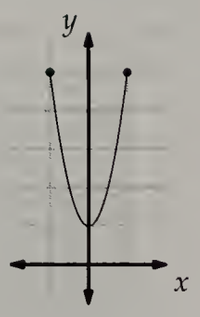
Solution for Problem 5.4: We can write $\cos 2\theta$ in terms of $\sin\theta$, so we can easily eliminate the parameter by substituting x for $\sin\theta$ in the resulting expression:

$$y = 3 - 2\cos 2\theta = 3 - 2(1 - 2\sin^2\theta) = 3 - 2(1 - 2x^2) = 4x^2 + 1.$$

What's wrong with this finish:

Bogus Solution: Therefore, the graph of the parametric equations is the parabola that results from graphing the quadratic $y = 4x^2 + 1$.

We see the problem when we look back at the parametric equations. We have $x = \sin\theta$, so x cannot be greater than 1 or less than −1. There is an angle θ such that $\sin\theta = x$ for each x from −1 to 1. Therefore, the graph of the parametric equations is the graph of $y = 4x^2 + 1$ for which $-1 \le x \le 1$, not the entire parabola that is produced by graphing $y = 4x^2 + 1$ without restriction on x. □



In Problem 5.4, we found that the parametric equations $x = \sin\theta$, $y = 3 - 2\cos 2\theta$ satisfy the equation $y = 4x^2 + 1$. However, the graph of the parametric equations is *not the same* as the graph of $y = 4x^2 + 1$. The fact that the parametric equations $x = \sin\theta$, $y = 3 - 2\cos 2\theta$ satisfy $y = 4x^2 + 1$ only tells us that the graph of the parametric equations is *part of* the graph of $y = 4x^2 + 1$. We then must consider the restrictions of the parametric equations to determine what portion of the graph of $y = 4x^2 + 1$ is the graph of the parametric equations.

WARNING!! As noted earlier, if parametric equations $x = f(t)$, $y = g(t)$ satisfy an equation in terms of x and y, then we only know that the graph of the parametric equations is *part of* the graph of the non-parametric equation. Sometimes the parametric equations have restrictions that limit the graph of the parametric equations to only a portion of the graph of the non-parametric equation.

Our first few examples have given us a sense of how parametric equations work, but not much sense for what parametric equations are good for. Unfortunately, you'll have to wait until your study of calculus and physics to

put parametric equations to heavy use, but Problem 5.3 gives some hint as to why parametric equations can be useful. In that problem, we showed that the equation $(x-3)^2+(y-4)^2=2^2$ can be modeled with the parametric equations $x=3+2\cos t$, $y=4+2\sin t$. We cannot write the equation $(x-3)^2+(y-4)^2=2^2$ such that one of x or y is a function of the other. However, our parametric equations allow us to write both x and y as functions of some other variable. The ability to express x and/or y as functions of another variable becomes particularly important in calculus, which offers very powerful tools that we can apply to functions.

Problem 5.5:

(a) Find parametric equations for x and y such that the graph of the parametric equations is the same as the graph of $x^2+y^2=1$.

(b) Is your answer to part (a) unique? Are there any other parametric equations that produce the same graph?

(c) Suppose a particle starts at the point $(0,-2)$, and moves at a constant speed clockwise around the origin along the graph of $x^2+y^2=4$. If the particle completes a full revolution in 4 seconds, then find parametric equations for the location of the particle in terms of t, where t is the time in seconds since the particle started moving.

Solution for Problem 5.5:

(a) The graph of $x^2+y^2=1$ is the unit circle. Any point on the unit circle is the terminal point of some angle θ. We defined the functions $\cos\theta$ and $\sin\theta$ such that the coordinates of this terminal point are $x=\cos\theta$, $y=\sin\theta$. Therefore, we say that $x=\cos\theta$, $y=\sin\theta$ is a **parameterization** of $x^2+y^2=1$ because the graph of the parametric equations $x=\cos\theta$, $y=\sin\theta$ is the same as the graph of $x^2+y^2=1$.

(b) Our answer in part (a) isn't the only parameterization of $x^2+y^2=1$. We also could have let $x=\cos(-\theta)$, $y=\sin(-\theta)$. Then, we still have $x^2+y^2=1$, and the parameterization produces every point on the circle. What's the difference between the two parameterizations $x=\cos\theta$, $y=\sin\theta$ and $x=\cos(-\theta)$, $y=\sin(-\theta)$? Both parameterizations produce the same curve, but if we care about how the curve is traced out as we vary θ, then there is a difference between these two parameterizations. As θ goes from 0 to 2π, the curve described by $x=\cos\theta$, $y=\sin\theta$ starts at $(1,0)$ and is produced by moving along the unit circle counterclockwise about the origin. On the other hand, as θ goes from 0 to 2π, the curve described by $x=\cos(-\theta)$, $y=\sin(-\theta)$ starts at $(1,0)$ and is generated *clockwise* from the origin.

There are infinitely many other ways we could have chosen the parametric equations. The equations $x=\sin 2t$, $y=\cos 2t$ work, as do $x=-\cos\pi t$, $y=-\sin\pi t$ and $x=\sin(t^2-t^3)$, $y=-\cos(t^2-t^3)$. While it doesn't matter which of these parameterizations we use to describe the graph of $x^2+y^2=1$, it can be very important which parameterization we choose when using parametric equations to model a process. In our next part, we'll see why.

(c) The first difference we see between this and part (a) is that the circle has radius 2 instead of radius 1. So, we might think to use the parameterization $x=2\cos t$, $y=2\sin t$. But, as described in part (b), a particle that follows the path given by $x=2\cos t$, $y=2\sin t$ will go *counterclockwise*, not clockwise. We saw in part (b) how to deal with that wrinkle. A particle that follows the path given by $x=2\cos(-t)$, $y=2\sin(-t)$ will travel clockwise as t increases. However, when $t=0$, this parameterization gives $(x,y)=(2,0)$, and we want the particle to start at $(0,-2)$. We take care of that with a phase shift, using the parameterization $x=2\cos\left(-t-\frac{\pi}{2}\right)$, $y=2\sin\left(-t-\frac{\pi}{2}\right)$.

We're not finished yet! We still have to worry about how fast the particle travels around the circle. The particle makes a full revolution in 4 seconds, so it must be back at $(0,-2)$ after 4 seconds. Therefore, we need the periods of the trig functions we use in our parameterization to be 4. Since the period of a function of the form $\sin at$ or $\cos at$ is $\frac{2\pi}{|a|}$ (for $a\neq 0$), we must have $|a|=\frac{\pi}{2}$. Therefore, a parameterization for the path of the particle is $x=2\cos\left(-\frac{\pi}{2}t-\frac{\pi}{2}\right)$, $y=2\sin\left(-\frac{\pi}{2}t-\frac{\pi}{2}\right)$.

This isn't the only way we can express this parameterization. We can use trigonometric identities to

simplify it. Since $\cos\left(\theta - \frac{\pi}{2}\right) = \sin\theta$ and $\sin\left(\theta - \frac{\pi}{2}\right) = -\cos\theta$, we can express our parameterization as $x = 2\sin\left(-\frac{\pi}{2}t\right)$, $y = -2\cos\left(-\frac{\pi}{2}t\right)$. Also, noting that $\sin(-\theta) = -\sin\theta$ and $\cos(-\theta) = \cos\theta$, we can express the parameterization as $x = -2\sin\left(\frac{\pi}{2}t\right)$, $y = -2\cos\left(\frac{\pi}{2}t\right)$.

We can check this answer by noting that when $t = 0$, we have $(x, y) = (0, -2)$, as expected. When $t = 1$, the particle has traveled one quarter-circle clockwise about the origin, and be at $(-2, 0)$. Sure enough, our parameterization gives $(x, y) = (-2, 0)$ when $t = 1$.

□

We say that we **parameterize an equation** when we find parametric equations for the variables in the equation. So, for example, we found that we can parameterize $x^2 + y^2 = 1$ as $x = \cos\theta$, $y = \sin\theta$. However, we also saw that this is not the only way to parameterize $x^2 + y^2 = 1$.

Concept: There are usually many ways to parameterize a given equation.

Problem 5.6: Parameterize the equation $x^2 - 3y^2 = 1$.

Solution for Problem 5.6: The graph of $x^2 - 3y^2 = 1$ is a hyperbola. What's wrong with this solution:

Bogus Solution: We add $3y^2$ to both sides and take the square root. This gives us the equation $x = \pm\sqrt{1 + 3y^2}$. So, the graph of $x = \pm\sqrt{1 + 3t^2}$, $y = t$ is the same as the graph of $x^2 - 3y^2 = 1$.

The problem here is that $\pm\sqrt{1 + 3t^2}$ is not a function. We can say that the graph of $x = \sqrt{1 + 3t^2}$, $y = t$ together with the graph of $x = -\sqrt{1 + 3t^2}$, $y = t$ gives us the graph of $x^2 - 3y^2 = 1$, but we still do not have a single pair of functions $f(t)$ and $g(t)$ such that the graph of $x = f(t)$, $y = g(t)$ is the graph of $x^2 - 3y^2 = 1$.

We've already seen a couple of examples in which trig identities helped us relate parametric equations to an equation in which x and/or y are squared, so we look for a way for apply trig identities here. Specifically, we look for identities in which we have a difference of squares of trig functions, since the equation $x^2 - 3y^2 = 1$ has a difference of squares of variable terms x and $\sqrt{3}y$. We might consider the trig identity

$$\cos^2\theta - \sin^2\theta = \cos 2\theta,$$

but that isn't promising for a couple of reasons. First, there's the $\cos 2\theta$ on the right—what will we do with that? Also, we figure that x probably can't be some constant times $\cos\theta$ or $\sin\theta$, since there are solutions to the equation $x^2 - 3y^2 = 1$ in which x and y are arbitrarily large numbers. Back to the drawing board.

An identity that involves squares of trig functions besides sine and cosine is $\tan^2\theta + 1 = \sec^2\theta$. Rearranging this gives $\sec^2\theta - \tan^2\theta = 1$. Aha! If we let $x^2 = \sec^2\theta$ and $3y^2 = \tan^2\theta$, then $x^2 - 3y^2$ becomes $\sec^2\theta - \tan^2\theta$, which we know equals 1. So, we can let $x = \sec\theta$ and $y = \frac{1}{\sqrt{3}}\tan\theta$ be the desired parametric equations. (We could have taken the negative square root in either equation and produced valid parametric equations, too.)

But why can we "take the square root" here to get a valid parameterization, but we couldn't do so in the Bogus Solution? We rejected the result of Bogus Solution, $x = \pm\sqrt{1 + 3t^2}$, $y = t$ because $\pm\sqrt{1 + 3t^2}$ is not a function. But what if took the positive square root and wrote the parameterization as $x = \sqrt{1 + 3t^2}$, $y = t$? The problem here is that this parameterization only produces points for which $x > 0$. All points with $x < 0$ that are on the graph of $x^2 - 3y^2 = 1$, such as $(-1, 0)$, are excluded from the graph of $x = \sqrt{1 + 3t^2}$, $y = t$.

With this in mind, we check to make sure that the graph of $x = \sec\theta$, $y = \frac{1}{\sqrt{3}}\tan\theta$ does capture the full hyperbola that is the graph of $x^2 - 3y^2 = 1$. As θ ranges from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$, we produce the branch of the hyperbola for which x is positive, where we produce the upper half of this branch when $0 \le \theta < \frac{\pi}{2}$ and the lower half when $-\frac{\pi}{2} < \theta \le 0$. Then, as θ ranges from $\frac{\pi}{2}$ to $\frac{3\pi}{2}$, we produce the branch of the hyperbola for which x is negative. So, the entire hyperbola is produced by graphing $x = \sec\theta$, $y = \frac{1}{\sqrt{3}}\tan\theta$. □

Problem 5.7: I'm riding my bike at a constant speed of 20 miles per hour on flat land when I ride over a dot of fresh yellow paint, thereby getting a yellow spot on my front wheel, which is 2 feet in diameter. I continue riding my bike in a straight line at 20 miles per hour. When t seconds have passed since running over the dot, how many feet is the yellow spot on my tire from the yellow dot I ran over?

Solution for Problem 5.7: What's wrong with this solution:

Bogus Solution: I ride 20 miles per hour for a time t, so the spot on the tire is now $20t$ away from the dot on the road.

There are several mistakes here. First, t is in seconds, my speed is in miles per hour, and we want the answer in feet. So, we have to convert 20 miles per hour to feet per second. Since there are 5280 feet in a mile, 20 miles is $20 \cdot 5280$ feet. There are 60 minutes in an hour and 60 seconds in each minute, so there are $60 \cdot 60 = 3600$ seconds in an hour. Therefore, $20 \cdot 5280$ feet per hour is $20 \cdot 5280/3600$ feet per second, which is approximately 29.33 feet per second.

The more interesting mistake comes from the fact that the answer is not simply $29.33t$ feet. The bike has moved $29.33t$ feet forward, because it has moved in a straight line at 29.33 feet per second for t seconds. However, the spot has not moved in a straight line. Sometimes it is touching the ground, sometimes it is at the top of the wheel, and most of the time it is somewhere in-between. The spot doesn't move in a straight line—as the center of the wheel moves forward, the spot rotates around it at a constant rate.

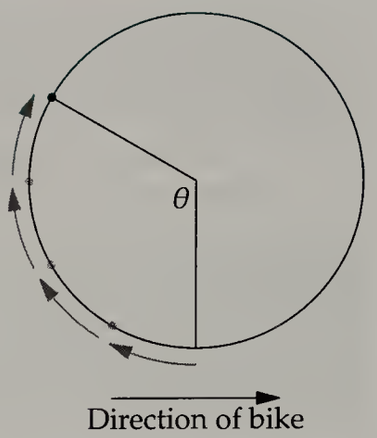
To figure out where the spot is after t seconds, we have to take both the forward motion of the bike and the rotary motion of the wheel into account. We'll do so by considering the horizontal motion of the spot and the vertical motion of the spot separately. We let $x(t)$ be the distance forward the spot has moved in t seconds, and $y(t)$ be how far off the ground the spot is after t seconds. In other words, we will find parametric equations for the location of the spot. Then, the distance from the starting point of the spot to its location after t seconds is simply $\sqrt{(x(t))^2 + (y(t))^2}$.

First, we tackle the horizontal movement. The center of the wheel moves forward at a speed of approximately $29.33t$ feet per second. However, the spot has not always moved the same distance forward as the center of the wheel has. Sometimes it's a little ahead of the center of the wheel, and sometimes it's a little behind the center of the wheel. To figure out exactly where the spot is relative to the center of the wheel, we have to figure out how much the spot has rotated around the center of the wheel. So, we have to figure out how much the wheel has turned.

For the moment, we'll forget about the fact that the wheel is moving forward and focus on the rotation, so we can figure out where the spot is relative to the center of the wheel at time t. The spot starts directly below the center of the wheel, and then turns an angle θ. We can figure out where the spot is relative to the center in much the same way we tackled the Ferris wheel problem on page 62. We think of the wheel as the unit circle. Moreover, the radius of the bike wheel is 1 foot, so, because we are using feet as our units, the wheel really is a unit circle.

Extra! The path the spot follows in Problem 5.7 is called a **cycloid**. Galileo gave the curve its name, and mathematicians of the 17th century argued so much about properties of the curve that it became known as the "Helen of Geometers." *Source: "Cycloid" by E. Weisstein*

The spot starts at the "bottom" of the circle, corresponding to an angle of $\frac{3\pi}{2}$. We let the wheel move in the "positive" direction, as shown, so that the wheel rotates clockwise. Suppose the spot rotates clockwise an angle θ about the center, as suggested by the grey spots and arrows. We can think of the spot as the terminal point of the angle $\frac{3\pi}{2} - \theta$. So, the spot's horizontal displacement from the wheel's center is $\cos\left(\frac{3\pi}{2} - \theta\right)$ feet and its vertical displacement from the center is $\sin\left(\frac{3\pi}{2} - \theta\right)$ feet. To find $x(t)$ and $y(t)$, we must combine the horizontal displacement relative to the center of the wheel with the fact that the wheel moves forward $29.33t$ feet, and we note that the center of the wheel is 1 foot off the ground. So, we have



$$x(t) \approx 29.33t + \cos\left(\frac{3\pi}{2} - \theta\right),$$
$$y(t) = 1 + \sin\left(\frac{3\pi}{2} - \theta\right).$$

Uh-oh. What do we do about θ? We have to express θ in terms of t, too. Fortunately, that's not too hard to do. When the wheel turns a full rotation of 2π, the bike moves forward by the circumference of the wheel, which is 2π feet, since the diameter of the wheel is 2 feet. So, the number of radians the wheel turns equals the distance the bike moves forwards in feet (this is a convenient coincidence caused by the fact that the wheel has radius 1 foot). Therefore, in the t seconds that the bike moves forward $29.33t$ feet, the wheel turns an angle $\theta = 29.33t$ radians. So, the parametric equations for the position of the spot are

$$x(t) \approx 29.33t + \cos\left(\frac{3\pi}{2} - 29.33t\right),$$
$$y(t) \approx 1 + \sin\left(\frac{3\pi}{2} - 29.33t\right).$$

Applying angle sum identities for sine and cosine, we have

$$x(t) \approx 29.33t - \sin 29.33t,$$
$$y(t) \approx 1 - \cos 29.33t.$$

Therefore, the distance from the spot on the wheel to the dot on the road is

$$\begin{aligned}\sqrt{(x(t))^2 + (y(t))^2} &\approx \sqrt{(29.33t - \sin 29.33t)^2 + (1 - \cos 29.33t)^2} \\ &\approx \sqrt{860.25t^2 - 58.66t \sin 29.33t + \sin^2 29.33t + 1 - 2\cos 29.33t + \cos^2 29.33t} \\ &= \sqrt{860.25t^2 - 58.66t \sin 29.33t - 2\cos 29.33t + 2}.\end{aligned}$$

□

Exercises

5.1.1 What is the slope of the line that is the graph of $x = 3t + 7$, $y = 8t - 5$?

5.1.2 Find the graph of the parametric equations $x = 3\cos 2t$, $y = -3\sin 2t$.

5.1.3 Find the graph of the parametric equations $x = -3 + 5\cos\theta$, $y = 2 + 5\sin\theta$.

5.1.4 Find the graph of the parametric equations $x = 4\cos 2t$, $y = 1 + \cos^2 2t$.

5.1.5 Find parametric equations for x and y such that $4x^2 + 9y^2 = 144$.

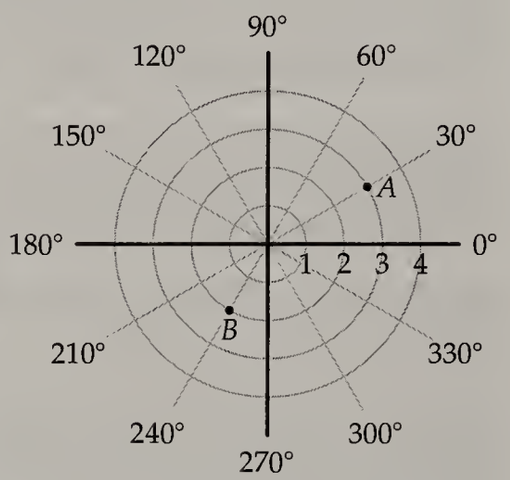
5.1.6 A particle starts at $(1,5)$ and travels at a constant speed in a circle around the point $(4,5)$. Find parametric equations to describe the location of the particle t seconds after it starts moving if the particle travels counterclockwise and takes 8π seconds to complete a full revolution around the circle.

5.1.7⋆ What is the graph of the parametric equations $x = \dfrac{4}{1+t^2}$, $y = \dfrac{4t}{1+t^2}$? **Hints:** 50, 128

5.2 Polar Coordinates

When we plot points on the Cartesian plane, we typically identify the location of each point with coordinates that denote where the point is horizontally and vertically with respect to a point we call the origin. These are the x- and y-coordinates that you're already used to; we sometimes refer to these as **rectangular coordinates**. While this is the most common method we use to identify the location of a point on a plane, it is not the only method.

Polar coordinates offer another method for denoting the location of a point on a plane. As with rectangular coordinates, we start with a point we call the origin and we identify the location of each point with an ordered pair, (r, θ). The r-coordinate is the distance from the point to the origin; we call this coordinate the **radial coordinate**. The θ-coordinate is the **angular coordinate**, which we define in the same way we related angles to points on the unit circle. A point with $\theta = 0°$ is directly to the right of the origin, a point with $\theta = 90°$ is directly above the origin, a point with $\theta = 180°$ is directly to the left of the origin, and so on, as depicted at right. For example, point A at right has polar coordinates $(3, 30°)$ and point B has polar coordinates $(2, 240°)$.



While each point only has one possible representation in rectangular coordinates, each point has multiple representations in polar coordinates. For example, we can add 360° to the angle, or subtract 360° from it. So, the polar coordinates of A could be written $(3, 390°)$ or $(3, 750°)$. We can also use a negative radius, which means "go in the direction opposite the one indicated by θ." For example, we could express A as $(-3, 210°)$, which tells us that A is 3 units in the direction opposite 210°, which is the direction of 30°. Similarly, the polar coordinates (r, θ) and $(-r, \theta + 180°)$ always refer to the same point. We can also use a negative angle, so point B above could be expressed as $(2, -120°)$.

WARNING!! Because we write both rectangular and polar coordinates as ordered pairs, it's easy to get confused about which we mean when we write a specific ordered pair. Ordered pairs of coordinates usually refer to rectangular coordinates, so if you're using polar coordinates, you should be very clear that you are doing so.

Problems

Problem 5.8: Convert each of the following rectangular coordinates to polar coordinates.

(a) $(0, -3)$

(b) $(5, 5)$

(c) $(-2\sqrt{3}, -6)$

Problem 5.9: Convert each of the following polar coordinates to rectangular coordinates.

(a) $(8, 60°)$

(b) $\left(12, -\frac{2\pi}{3}\right)$

(c) $(4\sqrt{2}, 315°)$

Problem 5.10:

(a) What are the rectangular coordinates of the point with polar coordinates $(1, \theta)$?

(b) What are the rectangular coordinates of the point with polar coordinates (r, θ)? (You can assume $r > 0$.)

(c) Suppose point P is (x, y) in rectangular coordinates and (r, θ) in polar coordinates, where $r > 0$. Find r in terms of x and y. How must we modify our answer if we remove the restriction $r > 0$? How is $\tan\theta$ related to x and y?

Problem 5.11: When we graph an equation that has the variables x and y, the graph consists of all points (x, y) on the Cartesian plane that satisfy the equation. Similarly, the graph of an equation that has the variables r and θ consists of all points with polar coordinates (r, θ) that satisfy the equation.

(a) What is the graph of $r = 5$?

(b) What is the graph of $\theta = 20°$?

Problem 5.12:

(a) Graph the equation $r = 6\sin\theta$. Use your graph to guess the rectangular form of this equation.

(b) Prove that $r = 6\sin\theta$ is equivalent to the rectangular form you find in part (a).

Problem 5.13: In rectangular form, the general form of an equation whose graph is the circle with center (h, k) and radius t is

$$(x-h)^2 + (y-k)^2 = t^2.$$

In this problem, we find the polar form of this equation. That is, we find the general form of an equation whose graph is the circle with center (r_0, α) (in polar coordinates) and radius t.

(a) Express h and k in terms of r_0 and α.

(b) Show that the polar form of the given rectangular form is $r^2 - 2rr_0\cos(\theta - \alpha) + r_0^2 = t^2$.

Problem 5.14:

(a) What are the rectangular coordinates of the point that results when $(5\sqrt{2}, 5\sqrt{6})$ is rotated $\frac{2\pi}{3}$ counterclockwise about the origin? **Hints:** 48

(b) Show that if the point (x, y) is rotated θ counterclockwise about the origin, the result has rectangular coordinates

$$(x\cos\theta - y\sin\theta, y\cos\theta + x\sin\theta).$$

Problem 5.15: Let h and k be constants, and let a and b be positive constants with $a > b$. The graph of an equation of the form

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$$

is an ellipse whose major axis has length $2a$. Every ellipse in the Cartesian plane with a horizontal major axis is the graph of an equation of this form.

The graph of the equation $7x^2 - 6\sqrt{3}xy + 13y^2 = 64$ is also an ellipse. Find the length of its major axis. **Hints:** 246

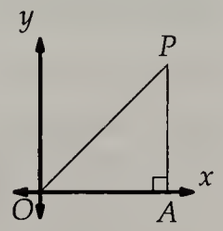
Problem 5.8: Convert each of the following rectangular coordinates to polar coordinates.

(a) $(0, -3)$ (b) $(5, 5)$ (c) $(-2\sqrt{3}, -6)$

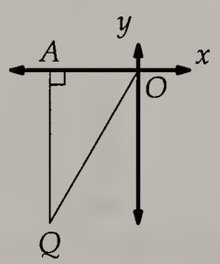
Solution for Problem 5.8:

(a) The point $(0, -3)$ is 3 units directly below the origin. Because the point is 3 units from the origin, the point has $r = 3$ in polar coordinates. Since the point is directly below the origin, the point has $\theta = 270°$. So, the polar coordinates of the point are $(3, 270°)$. As noted in the introduction to this section, there are many other ways we can express the polar coordinates of the point, such as $(3, -90°)$, $(-3, 90°)$, or $(3, 630°)$. Make sure you see why each of these describes the same point as $(3, 270°)$.

(b) The point $(5, 5)$ is point P at right. The rectangular coordinates give us $PA = AO = 5$, so $\triangle PAO$ is an isosceles right triangle. Therefore, we have $PO = 5\sqrt{2}$ and $\angle POA = 45°$, and the polar coordinates of point P are $(5\sqrt{2}, 45°)$.



(c) Point Q at right is $(-2\sqrt{3}, -6)$, so we have $AO = 2\sqrt{3}$ and $QA = 6$. To find the radius in polar coordinates, we note that $QO = \sqrt{QA^2 + AO^2} = 4\sqrt{3}$. Because the hypotenuse of $\triangle QAO$ is double one of the legs, we recognize that $\triangle QAO$ is a 30-60-90 right triangle. Since $\angle QOA$ is opposite the longer leg, we have $\angle QOA = 60°$. Point Q is 60° counterclockwise from the negative x-axis, so its angle in polar coordinates is $180° + 60° = 240°$. Therefore, point Q in polar coordinates is $(4\sqrt{3}, 240°)$.



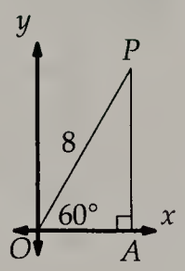
□

Problem 5.9: Convert each of the following polar coordinates to rectangular coordinates.

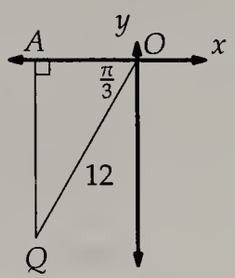
(a) $(8, 60°)$ (b) $\left(12, -\frac{2\pi}{3}\right)$ (c) $(4\sqrt{2}, 315°)$

Solution for Problem 5.9:

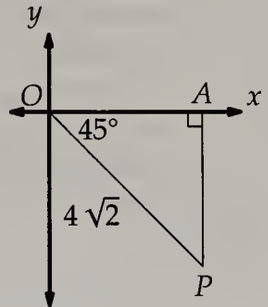
(a) Let P be the point with polar coordinates $(8, 60°)$, so that $PO = 8$ and $\angle POA = 60°$ in the diagram at right. From 30-60-90 triangle POA, we have $AO = \frac{PO}{2} = 4$ and $PA = AO\sqrt{3} = 4\sqrt{3}$, so the rectangular coordinates of point P are $(4, 4\sqrt{3})$.



(b) The point Q with polar coordinates $\left(12, -\frac{2\pi}{3}\right)$ is $\frac{2\pi}{3}$ clockwise from the positive x-axis. Therefore, Q is in the third quadrant, as shown at right. Drawing altitude $\overline{QA}$ produces right triangle QOA with hypotenuse $QO = 12$, from which we find $QA = 6\sqrt{3}$ and $OA = 6$. Since Q is in the third quadrant, both rectangular coordinates are negative, so the rectangular coordinates of Q are $(-6, -6\sqrt{3})$.



(c) A point with angle 315° in polar coordinates is in the fourth quadrant, 45° clockwise from the positive x-axis as shown at right. From 45-45-90 triangle PAO, we find $PA = AO = 4$, so the rectangular coordinates of P are $(4, -4)$. (Make sure you see why the y-coordinate is negative.)



□

Problem 5.10:

(a) What are the rectangular coordinates of the point with polar coordinates $(1, \theta)$?

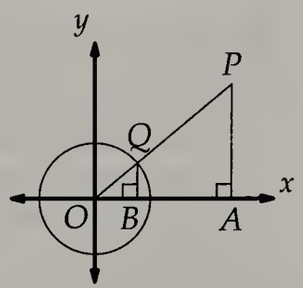
(b) What are the rectangular coordinates of the point with polar coordinates (r, θ)? (You can assume $r > 0$.)

(c) Suppose point P is (x, y) in rectangular coordinates and (r, θ) in polar coordinates, where $r > 0$. Find r in terms of x and y. How must we modify our answer if we remove the restriction $r > 0$? How is $\tan\theta$ related to x and y?

Solution for Problem 5.10:

(a) The point $(1, \theta)$ is one unit from the origin, so it is on the unit circle. This point results when we rotate the point $(1, 0)$ an angle of θ counterclockwise about the origin, which means that $(1, \theta)$ are the polar coordinates of the terminal point of θ. We used this point to define $\cos\theta$ and $\sin\theta$! Specifically, by our definitions of sine and cosine, the rectangular coordinates of the terminal point of θ are $(\cos\theta, \sin\theta)$.

(b) We can use our result from part (a) to quickly convert the polar coordinates (r, θ) to rectangular coordinates. In the diagram at right, point Q has polar coordinates $(1, \theta)$, so it is on the unit circle. Point P has polar coordinates (r, θ), so line $\overleftrightarrow{PQ}$ passes through the origin, as shown. Drawing altitudes from P and Q to the x-axis gives similar triangles PAO and QBO as shown. Because these triangles are similar, we have



$$\frac{PA}{QB} = \frac{AO}{BO} = \frac{PO}{QO}.$$

From the polar coordinates, we have $PO = r$ and $QO = 1$, so $\frac{PA}{QB} = \frac{PO}{QO}$ gives us $PA = r \cdot QB$, and $\frac{AO}{BO} = \frac{PO}{QO}$ gives us $AO = r \cdot BO$. Therefore, the distance from P to each axis is r times the corresponding distances from Q to each axis. Combining this with the fact that P is in the same quadrant as Q, we see that the rectangular coordinates of P are r times the rectangular coordinates of Q. So, the rectangular coordinates of P are $(r\cos\theta, r\sin\theta)$.

We can check this result by using it to confirm our answers to Problem 5.9. Converting $(8, 60°)$ to rectangular coordinates gives $(8\cos 60°, 8\sin 60°)$, which is $(4, 4\sqrt{3})$. Converting $\left(12, -\frac{2\pi}{3}\right)$ to rectangular coordinates gives $\left(12\cos\left(-\frac{2\pi}{3}\right), 12\sin\left(-\frac{2\pi}{3}\right)\right)$, which is $(-6, -6\sqrt{3})$. Converting $(4\sqrt{2}, 315°)$ to rectangular coordinates gives $(4\sqrt{2}\cos 315°, 4\sqrt{2}\sin 315°)$, which is $(4, -4)$. All three results match the ones we found earlier.

(c) The distance from (x, y) in rectangular coordinates to the origin is $\sqrt{x^2 + y^2}$. Therefore, when (x, y) is converted to polar coordinates such that $r > 0$, we have $r = \sqrt{x^2 + y^2}$. If we remove the restriction that $r > 0$, then we must instead write this relationship as $r^2 = x^2 + y^2$.

From the previous part, we have $x = r\cos\theta$ and $y = r\sin\theta$. If $x \neq 0$, we can divide the latter equation by the former to get $\frac{\sin\theta}{\cos\theta} = \frac{y}{x}$, so $\tan\theta = \frac{y}{x}$. If we restrict r to be nonnegative, then we *cannot* deduce from $\tan\theta = \frac{y}{x}$ that $\theta = \arctan\frac{y}{x}$, because the range of arctangent is $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. For example, if we have $(x, y) = (-3, -3)$, then $\theta = \arctan\frac{y}{x}$ gives us $\theta = \frac{\pi}{4}$, but the angle when we express $(-3, -3)$ in polar coordinates with $r > 0$ is $\frac{5\pi}{4}$ radians, which is not the same angle as $\frac{\pi}{4}$ radians. If we allow r to be negative, then we can let $\theta = \arctan\frac{y}{x}$, but then we must use the quadrant of (x, y) to determine whether r is positive or negative.

The best we can do to express θ simply in terms of x and y is to write $\tan\theta = \frac{y}{x}$ (for $x \neq 0$). If $x \neq 0$, there will be two values of θ with $0° \leq \theta < 360°$ that satisfy this equation. We choose whichever one corresponds to the quadrant of (x, y).

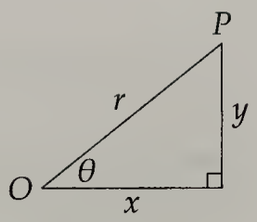
□

Important: If (x, y) in rectangular coordinates is the same point as (r, θ) in polar coordinates, then

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad r^2 = x^2 + y^2, \qquad \tan\theta = \frac{y}{x},$$

where $\tan\theta = \frac{y}{x}$ holds only if $x \neq 0$. If $x = 0$ and $y \neq 0$, then $\cos\theta = 0$ and $r^2 = y^2$. If $x = y = 0$, then $r = 0$ and θ can be anything.

There's no need to memorize these relationships; you can "see" them whenever you need them by visualizing the diagram at right, which depicts the relationship between a first quadrant point P and the origin O using both rectangular and polar coordinates. All of the equations in the box above can be read directly from the diagram.



We can use these relationships to give an algebraic explanation why the polar coordinates $(-r, \theta+180°)$ give the same point as (r, θ). Converting $(-r, \theta+180°)$ to rectangular coordinates, we have

$$x = (-r)\cos(\theta + 180°) = (-r)(-\cos\theta) = r\cos\theta,$$
$$y = (-r)\sin(\theta + 180°) = (-r)(-\sin\theta) = r\sin\theta,$$

which are the rectangular coordinates of (r, θ) as well.

Now that we understand how to locate points in polar coordinates, let's explore graphing equations in polar coordinates. When we graph an equation that has the variables x and y, the graph consists of all points (x, y) on the Cartesian plane that satisfy the equation. Similarly, the graph of an equation that has the variables r and θ consists of all points with polar coordinates (r, θ) that satisfy the equation.

Problem 5.11:

(a) What is the graph of $r = 5$?

(b) What is the graph of $\theta = 20°$?

Solution for Problem 5.11:

(a) The graph of $r = 5$ consists of all points that are 5 units away from the origin. Therefore, the graph is the circle with radius 5 centered at the origin.

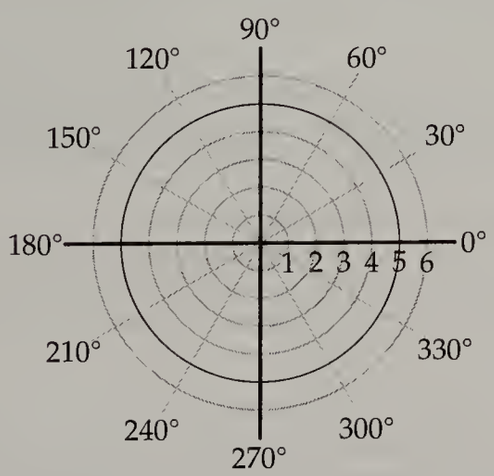


Figure 5.1: Graph of $r = 5$

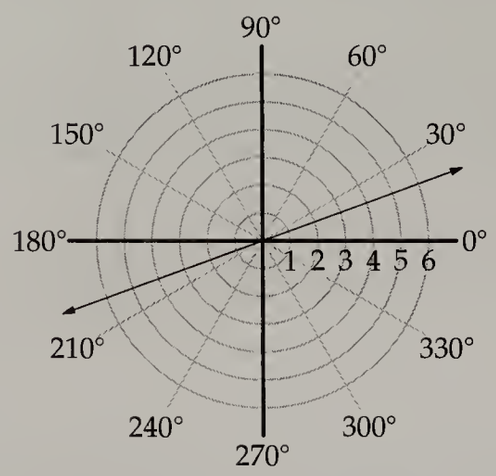


Figure 5.2: Graph of $\theta = 20°$

(b) Any point (r, θ) with $\theta = 20°$ and $r > 0$ in polar coordinates must be on the ray from the origin that makes a 20° angle with the positive x-axis such that the ray is in the first quadrant. Conversely, any point on this

ray can be expressed in polar coordinates with an angle of 20°. We must also include points for which the radius is 0 (the origin) or negative. If the radius r is negative, then the point $(r, 20°)$ is in the direction opposite the aforementioned ray. In other words, when r is negative, the point $(r, 20°)$ is on the line formed by continuing this ray past the origin into the third quadrant, as shown at right above. Every point on this ray in the third quadrant can be written as $(s, 200°)$ for some positive value of s, and such a point can also be written as $(-s, 200° - 180°)$, or $(-s, 20°)$. Therefore, the graph of $\theta = 20°$ is a line through the origin that passes through the first and third quadrants and makes an angle of 20° with the x-axis.

□

Having tackled the simplest equations, those in which r or θ is constant, let's look at an equation that includes both r and θ.

Problem 5.12:

(a) Graph the equation $r = 6\sin\theta$. Use your graph to guess the rectangular form of this equation.

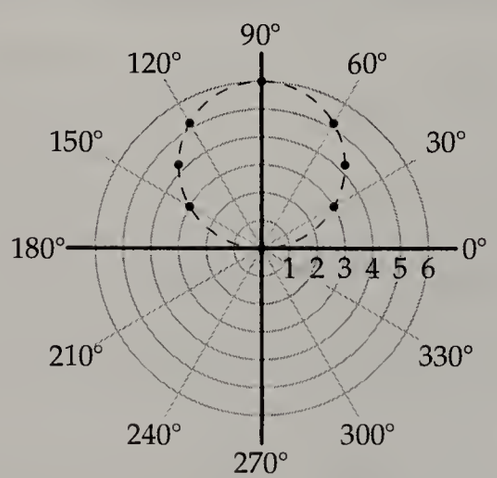
(b) Prove that $r = 6\sin\theta$ is equivalent to the rectangular form you find in part (a).

Solution for Problem 5.12:

(a) We get a feel for the graph by finding r for several values of θ. When $\theta = 0$, we have $r = 0$, so the graph passes through the origin. When $0° < \theta < 180°$, we have $r > 0$, so there are points in the first and second quadrants in the graph, corresponding to $0° < \theta < 90°$ and $90° < \theta < 180°$, respectively. The largest value of r occurs when $\sin\theta = 1$, which is when $\theta = 90°$. This gives $r = 6$, so the point 6 units directly above the origin is on the graph. Since $6\sin\theta$ increases from 0 to 6 as θ goes from 0° to 90°, and then goes back down from 6 to 0 as θ goes from 90° to 180°, the graph goes from the origin through the first quadrant to the point 6 above the origin ((0, 6) in rectangular coordinates, $(6, 90°)$ in polar coordinates), and then back down through the second quadrant to the origin.

When $180° < \theta < 360°$, the value of $6\sin\theta$ is negative. As we pick a few values of θ in this range, we notice that each produces a point we already found when examining $0° < \theta < 180°$, since a negative radius means the point is in the direction opposite θ. For example, if $\theta = 210°$, we have $r = 6\sin\theta = -3$, which gives us the point $(-3, 210°)$. But this point is the same as $(3, 210° - 180°)$, which is $(3, 30°)$, a point we already know is on the graph. Similarly, since $\sin(\alpha + 180°) = -\sin\alpha$, we can write the point $(6\sin(\alpha + 180°), \alpha + 180°)$ as $(-6\sin\alpha, \alpha + 180°)$, which is the same as $(6\sin\alpha, \alpha)$. Therefore, we only have to worry about the angles $0° \le \theta < 180°$ when constructing the graph of $r = 6\sin\theta$. Our table and the corresponding points are below.

θ	(r, θ)
0°	$(0, 0°)$
30°	$(3, 30°)$
45°	$(3\sqrt{2}, 45°)$
60°	$(3\sqrt{3}, 60°)$
90°	$(6, 90°)$
120°	$(3\sqrt{3}, 120°)$
135°	$(3\sqrt{2}, 135°)$
150°	$(3, 150°)$



These points suggest that our graph is the circle with center (0, 3) and radius 3. We know that in rectangular coordinates, an equation of the form

$$(x - h)^2 + (y - k)^2 = t^2$$

has a graph that is a circle with center (h,k) and radius t, so we guess that the graph of the equation $r = 6\sin\theta$ is the same as the graph of the equation $x^2 + (y-3)^2 = 3^2$.

(b) Now that we have some clue what we're looking for, we know just how to look for it. We wish to show that the graphs of $r = 6\sin\theta$ and $x^2 + (y-3)^2 = 3^2$ are the same. We start by looking at the equations we found in Problem 5.10 to relate polar and rectangular coordinates:

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad r^2 = x^2 + y^2, \qquad \tan\theta = \frac{y}{x}.$$

Our goal is to use these relationships to convert $r = 6\sin\theta$ to $x^2 + (y-3)^2 = 3^2$ (or vice versa). First, we note that the origin is in both graphs. For all other points, we have $r > 0$, so we don't have to worry about inadvertently multiplying or dividing by 0 when we multiply or divide an equation by r. Here are a couple ways that we can convert the polar coordinate equation $r = 6\sin\theta$ to the rectangular coordinate equation $x^2 + (y-3)^2 = 3^2$:

Method 1: Substitution for $\sin\theta$. We write $y = r\sin\theta$ as $\sin\theta = \frac{y}{r}$. Substituting this into $r = 6\sin\theta$ gives $r = \frac{6y}{r}$. Multiplying both sides by r gives $r^2 = 6y$. Aha! We have an equation for r^2 in terms of x and y. Substituting $r^2 = x^2 + y^2$ gives $x^2 + y^2 = 6y$. Rearranging this gives $x^2 + y^2 - 6y = 0$, and completing the square in y gives $x^2 + (y-3)^2 = 3^2$.

Method 2: Manipulation to introduce r^2 *and* $r\sin\theta$. We know that we can convert $r\sin\theta$ into y, and convert r^2 into $x^2 + y^2$, so we multiply both sides of $r = 6\sin\theta$ by r to give $r^2 = 6r\sin\theta$. We can now write this equation as $x^2 + y^2 = 6y$, and finish as before. The key step in this second approach is trying to force the expressions $r\sin\theta$ and r^2 into the equation, since we can then eliminate the polar variables by substituting the equivalent rectangular variable expressions.

We also could have started with the rectangular coordinate equation $x^2 + (y-3)^2 = 3^2$ and converted it to the polar coordinate equation $r = 6\sin\theta$. We start by expanding and rearranging $x^2 + (y-3)^2 = 3^2$ to get $x^2 + y^2 = 6y$. Then, for any point (x,y), we can let $r^2 = x^2 + y^2$ and $y = r\sin\theta$, and we have $r^2 = 6r\sin\theta$. Dividing $r^2 = 6r\sin\theta$ by r gives $r = 6\sin\theta$.

We conclude that the graph of $r = 6\sin\theta$ is the same as the graph of $x^2 + (y-3)^2 = 3^2$, which is a circle with radius 3 centered at the point with rectangular coordinates $(0,3)$.

□

Problem 5.13: In rectangular form, the general form of an equation whose graph is the circle with center (h,k) and radius t is

$$(x-h)^2 + (y-k)^2 = t^2.$$

Find the polar form of this equation. That is, find the general form of an equation whose graph is the circle with center (r_0,α) (in polar coordinates) and radius t.

Solution for Problem 5.13: Our goal is to turn the equation in terms of x, y, h, and k into an equation in terms of r, θ, r_0, and α. Again, our equations from Problem 5.10 are the key. Here, the conversion is a bit easier than in Problem 5.12. Since we are given the rectangular form, we can simply substitute $x = r\cos\theta$ and $y = r\sin\theta$ to replace the rectangular variables (x,y) with the polar variables (r,θ). This gives us

$$(r\cos\theta - h)^2 + (r\sin\theta - k)^2 = t^2.$$

However, we still have to get rid of h and k—these are rectangular coordinates as well. Just as we can use $x = r\cos\theta$ and $y = r\sin\theta$ to convert the variables (x,y) to the variables (r,θ), we can use them to express the center's rectangular coordinates (h,k) in terms of its polar coordinates (r_0,α). We therefore have $h = r_0\cos\alpha$ and $k = r_0\sin\alpha$, and the equation of our circle is

$$(r\cos\theta - r_0\cos\alpha)^2 + (r\sin\theta - r_0\sin\alpha)^2 = t^2.$$

Expanding the squares gives

$$r^2\cos^2\theta - 2rr_0\cos\theta\cos\alpha + r_0^2\cos^2\alpha + r^2\sin^2\theta - 2rr_0\sin\theta\sin\alpha + r_0^2\sin^2\alpha = t^2.$$

Seeing squares of cosine and sine makes us look for ways to add them to get 1. Fortunately, factoring does the trick both for both θ and α. We can rewrite the equation as

$$r^2(\cos^2\theta + \sin^2\theta) - 2rr_0(\cos\theta\cos\alpha + \sin\theta\sin\alpha) + r_0^2(\cos^2\alpha + \sin^2\alpha) = t^2.$$

Applying trigonometric identities to each of the three expressions in parentheses, the equation becomes

$$r^2 - 2rr_0\cos(\theta - \alpha) + r_0^2 = t^2.$$

□

Problem 5.14:

(a) What are the rectangular coordinates of the point that results when $(5\sqrt{2}, 5\sqrt{6})$ is rotated $\frac{2\pi}{3}$ counterclockwise about the origin?

(b) Show that if the point (x, y) is rotated θ counterclockwise about the origin, the result has rectangular coordinates $(x\cos\theta - y\sin\theta, y\cos\theta + x\sin\theta)$.

Solution for Problem 5.14:

(a) Aside from the name of this section, we have a number of clues to try polar coordinates. First, we don't have good tools for handling rotations in rectangular coordinates (aside from rotating by multiples of $\frac{\pi}{2}$). Second, rotation has to do with angles, and we have angles in polar coordinates, not in rectangular coordinates. So, we start by converting $(5\sqrt{2}, 5\sqrt{6})$ to polar coordinates. We have $r^2 = x^2 + y^2 = 50 + 150 = 200$, so we can let $r = 10\sqrt{2}$. We also have $\tan\theta = \frac{y}{x} = \sqrt{3}$. Since the point is in the first quadrant and we are using the positive value of r, we have $\theta = \frac{\pi}{3}$. Therefore, the point $(5\sqrt{2}, 5\sqrt{6})$ in rectangular coordinates is the same as the point $\left(10\sqrt{2}, \frac{\pi}{3}\right)$ in polar coordinates.

Finding the result of rotating $\left(10\sqrt{2}, \frac{\pi}{3}\right)$ counterclockwise $\frac{2\pi}{3}$ about the origin is easy. We simply add $\frac{2\pi}{3}$ to the angle, since rotation about the origin changes the angle by the amount of the rotation, but it doesn't change the distance from the point to the origin. So, rotating $\left(10\sqrt{2}, \frac{\pi}{3}\right)$ counterclockwise $\frac{2\pi}{3}$ gives us $(10\sqrt{2}, \pi)$. Converting this back to rectangular coordinates gives $(-10\sqrt{2}, 0)$.

(b) We use our first part as a guide.

Concept: When proving a formula, it's often useful to work out a specific example, and then use the example as a guide to find the proof.

Let (r, α) be the polar coordinates of the point with rectangular coordinates (x, y). Rotating (r, α) counterclockwise about the origin by θ gives the point with polar coordinates $(r, \alpha + \theta)$. The rectangular coordinates of this point are $(r\cos(\alpha + \theta), r\sin(\alpha + \theta))$. Applying angle sum identities, we can write these coordinates as

$$(r\cos\alpha\cos\theta - r\sin\alpha\sin\theta, r\sin\alpha\cos\theta + r\sin\theta\cos\alpha).$$

That's not quite what we were looking for—the result we wish to prove has x and y, and no r or α. Since (x, y) in rectangular coordinates is (r, α) in polar coordinates, we have $x = r\cos\alpha$ and $y = r\sin\alpha$. We can substitute x for $r\cos\alpha$ and y for $\sin\alpha$ in the long coordinates above to give the simpler-looking

$$(x\cos\theta - y\sin\theta, y\cos\theta + x\sin\theta).$$

□

Important: The image of the point (x, y) upon a rotation θ counterclockwise about the origin is $(x\cos\theta - y\sin\theta, y\cos\theta + x\sin\theta)$.

Problem 5.15: Let h and k be constants, and let a and b be positive constants with $a > b$. The graph of an equation of the form

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$$

is an ellipse whose major axis has length $2a$. Every ellipse in the Cartesian plane with a horizontal major axis is the graph of an equation of this form.

The graph of the equation $7x^2 - 6\sqrt{3}xy + 13y^2 = 64$ is also an ellipse. Find the length of its major axis.

Solution for Problem 5.15: Let $\mathcal{E}$ be the graph of $7x^2-6\sqrt{3}xy+13y^2 = 64$. If we can manipulate $7x^2-6\sqrt{3}xy+13y^2 = 64$ into the form $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$, then we can easily find the length of the major axis. Unfortunately, the xy term is a problem. There is no xy term in the form we're given for an ellipse with a horizontal major axis, so $\mathcal{E}$ must be an ellipse whose major axis is not horizontal. If our ellipse had a horizontal major axis, we could solve the problem quickly. So, we try to relate $\mathcal{E}$ to an ellipse with a horizontal major axis that is the same length as the major axis of $\mathcal{E}$.

Rotation does not affect length, so we can view $\mathcal{E}$ as the result of rotating an ellipse $\mathcal{E}'$ with a horizontal major axis. Suppose the required rotation is a counterclockwise rotation by θ, so that for every point (x', y') on $\mathcal{E}'$, there is a point (x, y) on $\mathcal{E}$ that is a counterclockwise rotation by θ of (x', y'). From the result of Problem 5.14, we have

$$x = x'\cos\theta - y'\sin\theta,$$
$$y = x'\sin\theta + y\cos\theta.$$

Substituting these into $7x^2 - 6\sqrt{3}xy + 13y^2 = 64$ produces

$$7(x'\cos\theta - y'\sin\theta)^2 - 6\sqrt{3}(x'\cos\theta - y'\sin\theta)(x'\sin\theta + y'\cos\theta) + 13(x'\sin\theta + y'\cos\theta)^2 = 64. \quad (5.1)$$

Yikes. Multiplying that out looks scary. Fortunately, we don't have to do so just yet. Our goal is to eliminate the $x'y'$ term. So, we just look at the $x'y'$ terms in the expansion of the left side, which are

$$-14x'y'\cos\theta\sin\theta - 6\sqrt{3}x'y'\cos^2\theta + 6\sqrt{3}x'y'\sin^2\theta + 26x'y'\cos\theta\sin\theta.$$

We want these terms to cancel, so we must have

$$-14\cos\theta\sin\theta - 6\sqrt{3}\cos^2\theta + 6\sqrt{3}\sin^2\theta + 26\cos\theta\sin\theta = 0.$$

Simplifying a bit gives us $12\cos\theta\sin\theta - 6\sqrt{3}(\cos^2\theta - \sin^2\theta) = 0$. Now we see a chance to use trig identities. We have $\cos\theta\sin\theta = \frac{1}{2}\sin 2\theta$ and $\cos^2\theta - \sin^2\theta = \cos 2\theta$, so we can write the equation as

$$6\sin 2\theta - 6\sqrt{3}\cos 2\theta = 0,$$

from which we have $(\sin 2\theta)/(\cos 2\theta) = \sqrt{3}$, or $\tan 2\theta = \sqrt{3}$. Since $\tan 60° = \sqrt{3}$, we find that $2\theta = 60°$, so $\theta = 30°$. This means that $\mathcal{E}$ is a 30° counterclockwise rotation of $\mathcal{E}'$ about the origin.

We can now find the equation whose graph is $\mathcal{E}'$. Substituting $\theta = 30°$ into Equation (5.1) gives

$$7\left(x'\cdot\frac{\sqrt{3}}{2} - y'\cdot\frac{1}{2}\right)^2 - 6\sqrt{3}\left(x'\cdot\frac{\sqrt{3}}{2} - y'\cdot\frac{1}{2}\right)\left(x'\cdot\frac{1}{2} + y'\cdot\frac{\sqrt{3}}{2}\right) + 13\left(x'\cdot\frac{1}{2} + y'\cdot\frac{\sqrt{3}}{2}\right)^2 = 64.$$

Multiplying both sides by 4 gets rid of the fractions and gives

$$7(\sqrt{3}x' - y')^2 - 6\sqrt{3}(\sqrt{3}x' - y')(x' + \sqrt{3}y') + 13(x' + \sqrt{3}y')^2 = 256.$$

Expanding and then simplifying the left side gives

$$16(x')^2 + 64(y')^2 = 256.$$

The $x'y'$ terms all canceled out, as expected. Dividing both sides by 256 puts the equation in the form of an ellipse described in the problem,

$$\frac{(x')^2}{4^2} + \frac{(y')^2}{2^2} = 1.$$

Comparing this to

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

we find $a = 4$, so the length of the major axis is $2a = 8$. □

Exercises

5.2.1 Convert each of the following rectangular coordinates to polar coordinates. (You may use a calculator if necessary.)

(a) $(-5, 0)$

(b) $(6\sqrt{2}, 6)$

(c) $(-4, -4\sqrt{3})$

5.2.2 Convert each of the following polar coordinates to rectangular coordinates.

(a) $(7, 210°)$

(b) $(9, -180°)$

(c) $\left(-8, \frac{9\pi}{4}\right)$

5.2.3 Graph the equation $r = \dfrac{4}{2\cos\theta - 5\sin\theta}$.

5.2.4 The final form we found in Problem 5.13 for an equation whose graph is a circle does not match the form of the equation we graphed in Problem 5.12. But the graph we found in Problem 5.12 is indeed a circle. This suggests we can write the equation we graphed in Problem 5.12, which is $r = 6\sin\theta$, in the form we found in Problem 5.13, which is $r^2 - 2rr_0\cos(\theta - \alpha) + r_0^2 = t^2$. Show that we can do so.

5.2.5 Let A and B be points with polar coordinates $(6, 137°)$ and $(12, \theta)$, respectively, and let O be the origin.

(a) Find all possible values of θ with $0° \le \theta < 360°$ such that $\overline{AO} \perp \overline{BO}$.

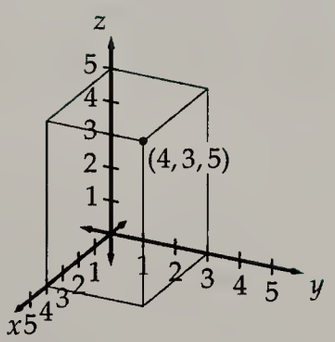
(b) Find all possible values of θ with $0° \le \theta < 360°$ such that $\overline{AB} \perp \overline{AO}$.

5.2.6 What does the graph of $r = 2 + 3\theta$ look like, where $\theta \ge 0$ and we express θ in radians when computing r?

5.2.7 A regular hexagon with center at the origin has one vertex at $(3, 3)$. Find all vertices of the hexagon that are in the second quadrant.

5.3 Coordinates in Three Dimensions

Just as we use the horizontal and vertical coordinates (x, y) to identify the location of a point in the two-dimensional Cartesian plane, we can identify the location of a point in three-dimensional **Cartesian space** with three coordinates (x, y, z). We call the plane in which $z = 0$ the ***xy*-plane**. Points on this plane are identified by the coordinates $(x, y, 0)$, where the x and y have the horizontal and vertical meanings they have in the Cartesian plane. For other points (x, y, z) in space, the z-coordinate indicates how far above ($z > 0$) or below ($z < 0$) the xy-plane the point is. The x- and y-coordinates of (x, y, z) are the same as those of the point in the $z = 0$ plane that is directly above or below the point (x, y, z).



As shown in the diagram at right, we have a third axis in three dimensions, which we call the z-axis. In this diagram we depict the point $(4, 3, 5)$. We include edges of a right rectangular prism with opposite vertices at $(0, 0, 0)$ and $(4, 3, 5)$ to help show the location of $(4, 3, 5)$ in space.

Just as all points (x, y, z) with $z = 0$ form the xy-plane, the ***yz*-plane** consists of all points of the form $(0, y, z)$, and the ***xz*-plane** consists of all points of the form form $(x, 0, z)$.

In most of the three-dimensional diagrams we include in this book, we'll omit the negative portions of each axis, and remove the tick marks and numbers along the axes, since these diagrams can get very cluttered if all of this information is included. Sometimes, we'll even exclude the axes, and just include the key geometric figures.

Problems

Problem 5.16: We start our exploration of coordinates in three dimensions by finding a formula for the distance between two points in space.

(a) Let A be $(-3, 4, 6)$ and B be $(4, -1, 2)$. Find AB.

(b) Find a formula for the distance between the points (x_1, y_1, z_1) and (x_2, y_2, z_2).

Problem 5.17: Describe the graph of the equation $x^2 + y^2 + z^2 - 6x - 4y + 8z = 7$.

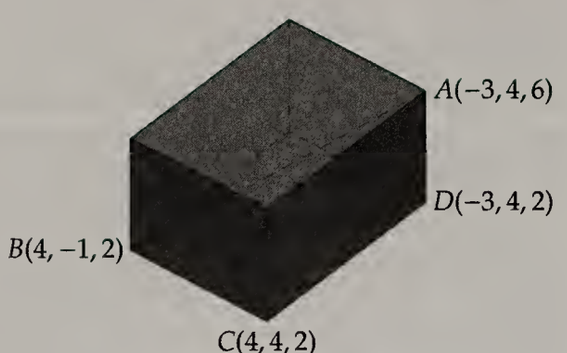
Problem 5.16:

(a) Let A be $(-3, 4, 6)$ and B be $(4, -1, 2)$. Find AB.

(b) Find a formula for the distance between the points (x_1, y_1, z_1) and (x_2, y_2, z_2).

Solution for Problem 5.16:

(a) We find AB by considering the rectangular prism with A and B as opposite vertices and with each edge parallel to a coordinate axis, as shown at right. We can then find AB by applying the Pythagorean Theorem twice. As shown in the diagram, the points $(4, 4, 2)$ and $(-3, 4, 2)$ are vertices of the prism. Let these be C and D, respectively. Right triangle BCD gives us

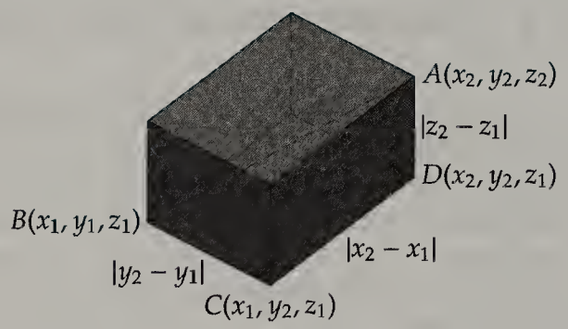


$$BD = \sqrt{BC^2 + CD^2} = \sqrt{25 + 49} = \sqrt{74}.$$

Then, right triangle ADB gives us

$$AB = \sqrt{BD^2 + AD^2} = \sqrt{74 + 16} = 3\sqrt{10}.$$

(b) We follow the same strategy as in part (a). We construct the rectangular prism with (x_1, y_1, z_1) and (x_2, y_2, z_2) as opposite vertices and with edges parallel to the coordinate axes, as shown. From right triangle BCD, we have $BD^2 = CD^2 + BC^2$, and from right triangle ABD, we have



$$AB^2 = BD^2 + AD^2 = CD^2 + BC^2 + AD^2,$$

so $AB = \sqrt{CD^2 + BC^2 + AD^2}$. Writing CD, BC, and AD in terms of the coordinates as shown, we see that the distance between (x_1, y_1, z_1) and (x_2, y_2, z_2) is

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

□

Important: The distance between the points (x_1, y_1, z_1) and (x_2, y_2, z_2) is

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

Problem 5.17: Describe the graph of the equation $x^2 + y^2 + z^2 - 6x - 4y + 8z = 7$.

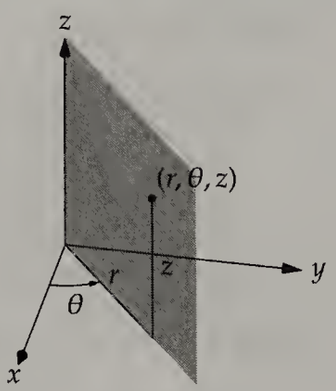
Solution for Problem 5.17: The equation strongly resembles a two-variable equation whose graph is a circle, so we suspect the graph is a sphere. When confronted with similar equations with two variables, we typically complete the squares in the quadratics, so we try that here. Completing the square in x, in y, and in z gives us

$$(x - 3)^2 + (y - 2)^2 + (z + 4)^2 = 7 + 9 + 4 + 16 = 36.$$

The left side looks a lot like the distance formula in three dimensions. In fact, the expression on the left equals the square of the distance between (x, y, z) and $(3, 2, -4)$. Since the right side is 6^2, the equation tells us that the distance between (x, y, z) and $(3, 2, -4)$ is 6. Sure enough, the graph of all such points is a sphere with center $(3, 2, -4)$ and radius 6. □

Just as Cartesian coordinates are not the only way to describe the location of a point in a plane, the **rectangular coordinates** (x, y, z) described above are not the only way to describe the location of a point in space. We'll now introduce two other ways to describe the location of a point in space.

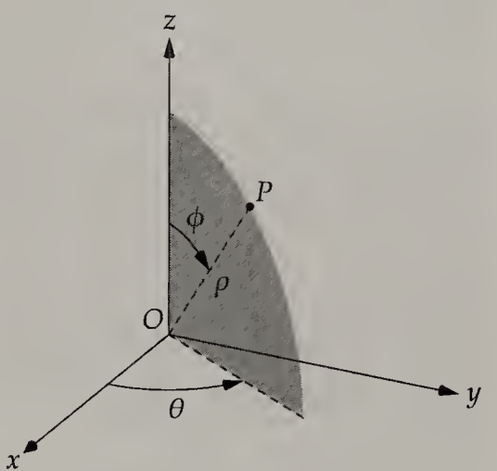
The first is **cylindrical coordinates**. In cylindrical coordinates, the x- and y-coordinates are replaced by r- and θ-coordinates such that the point in the Cartesian plane with rectangular coordinates (x, y) has polar coordinates (r, θ). The third cylindrical coordinate is the same z-coordinate as in rectangular coordinates. A depiction of the point (r, θ, z) in cylindrical coordinates is shown at right. When expressing a point in polar coordinates, the radius always comes first, the angle second, and the z-coordinate last.



Relating the cylindrical r- and θ-cylindrical coordinates to rectangular coordinates in three dimensions is the same as relating polar coordinates (r, θ) to rectangular coordinates (x, y) in two dimension. If the cylindrical coordinates (r, θ, z) represent the same point as the rectangular coordinates (x, y, z), then $x = r\cos\theta$ and $y = r\sin\theta$. As with polar coordinates, we usually choose r to be nonzero and θ such that $0 \le \theta < 2\pi$. However, r can be negative and θ can be outside the interval $[0, 2\pi)$, just as with polar coordinates.

Cylindrical coordinates aren't a natural analogue to polar coordinates. When a point is expressed in polar coordinates, the absolute value of the r-coordinate is the distance from the point to the origin, but in cylindrical coordinates, the r-coordinate gives the distance from the point to the z-axis, not the origin. A more natural analogue in three dimensions to the polar coordinates of two dimensions starts with letting the first coordinate represent the distance from the origin to the point. We do just that when we describe a point's location with **spherical coordinates**.

The first spherical coordinate is ρ, pronounced "roh," which is the distance from the point to the origin. But specifying the distance from the origin to a point is not enough to determine exactly where the point is. In order to describe the location of the point, we need a pair of angles as well. The first of the two angles we use to describe a point P in spherical coordinates is the same as the angle in cylindrical coordinates. We think of the half-plane that starts from the z-axis and extends through P. The angle θ is the angle between this plane and the positive x-axis, as shown, like the angle θ in cylindrical coordinates. We define the second angle ϕ, pronounced "fee," as the angle between $\overline{OP}$ and the positive z-axis, as shown. Therefore, every point in space can be described with a triple (ρ, θ, ϕ), with $\rho \geq 0$, $0 \leq \theta < 2\pi$ and $0 \leq \phi \leq \pi$.



While we usually restrict ρ, θ, and ϕ such that $\rho \geq 0$, $0 \leq \theta < 2\pi$ and $0 \leq \phi \leq \pi$, we don't have to. Allowing ρ to be negative has the same effect as allowing r to be negative in polar coordinates. When r is negative in polar coordinates, the point is in the direction from the origin that is opposite the direction described by the θ-coordinate. When ρ is negative in polar coordinates, the point is in the direction from the origin that is opposite the direction described by the θ- and ϕ-coordinates. So, for example, the point $(1, 0, 1)$ can be described as both $\left(\sqrt{2}, 0, \frac{\pi}{4}\right)$ and $\left(-\sqrt{2}, \pi, \frac{3\pi}{4}\right)$ in spherical coordinates.

WARNING!! When reading about spherical coordinates in another source, make sure you know what each variable stands for and what the restrictions on each are. While most sources will define the two angles in spherical coordinates in essentially the same way we have here, some will use different variables than we have. Even more confusing, some will use the same variables, but reverse their meanings. Furthermore, some will restrict ρ to be nonnegative, and some will use r rather than ρ.

Problems

Problem 5.18:

(a) Describe the graph of all points with cylindrical coordinates (r, θ, z) such that $r = 7$. Express the equation in rectangular coordinates.

(b) Describe the graph of $\theta = \frac{\pi}{3}$ in cylindrical coordinates. Express the equation in rectangular coordinates.

Problem 5.19: Suppose (ρ, θ, ϕ) are the spherical coordinates of point P, where $\rho \geq 0$, and that (x, y, z) are the rectangular coordinates of point P.

(a) Find ρ in terms of x, y, and z.

(b) Let O be the origin and A be the foot of the altitude from P to the z-axis. Find AO in terms of ρ, θ, and/or ϕ.

(c) Find x, y, and z in terms of ρ, θ, and ϕ.

Problem 5.20:
(a) Convert the cylindrical coordinates $\left(4, \frac{\pi}{3}, 4\right)$ to rectangular and spherical coordinates.

(b) Convert the spherical coordinates $\left(7, \frac{7\pi}{4}, \frac{\pi}{3}\right)$ to rectangular and cylindrical coordinates.

Problem 5.21: Convert each of the following rectangular coordinates to cylindrical coordinates and to spherical coordinates.

(a) $(6, -6, -2\sqrt{6})$ (b) $(-6, 0, 6\sqrt{3})$

Problem 5.22:
(a) Describe the graph of the equation $z^2 = x^2 + y^2$.

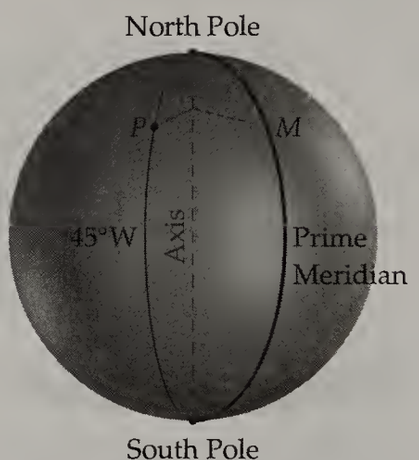
(b) Find parametric equations for x, y, and z for the equation $z^2 = x^2 + y^2$. (You will need *two* parameters.)

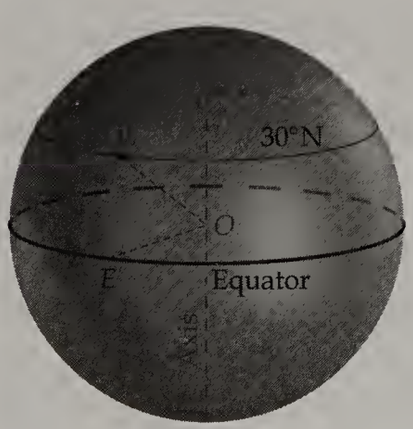
Problem 5.23:
(a) Describe the graph of all points with spherical coordinates (ρ, θ, ϕ) such that $\rho = 5$. Express the equation in rectangular coordinates, and give a parameterization of the equation for x, y, and z.

(b) Describe the graph of $\phi = \frac{2\pi}{3}$ in spherical coordinates. Express the equation in rectangular coordinates.

For the final problem in this section, we need to understand how we determine the **latitude** and **longitude** that we typically use to locate points on the Earth's surface. To do so, we start with a **great circle** of a sphere, which is a cross-section of the sphere that has the same center as the sphere. We choose the great circle that is in the plane perpendicular to the axis of the Earth to be the **equator** of the earth. The equator has latitude 0°, and it divides the Earth into the Northern hemisphere and the Southern hemisphere.

The North and South Poles of the Earth are the points at which the axis intersects the surface of the Earth. To define longitude, we first choose one semicircle along the surface of the Earth, with the poles as endpoints, to be the **prime meridian**. We say that the longitude of any point on the prime meridian is 0°. For any point P not on the prime meridian, we determine its longitude by considering the plane through P parallel to the equator. Let this plane intersect the axis at A and the prime meridian at M. Then, $\angle PAM$ is the longitude of P. We have one little wrinkle here—we add a direction to the longitude to tell us which way to go, east or west, to get to P fastest from M. So, if $\angle PAM$ is 45° and we go west from M to get to P, then the latitude of P is 45°W, where the W stands for "west."





Now that we have longitude, we can use it to define latitude. First, we say the North and South Poles have latitudes 90°N and 90°S, respectively. For any other point T not on the equator, we let E be the point on the equator with the same longitude as T and let O be the center of the Earth. Then, $\angle TOE$ gives us the latitude, where we assign the latitude the designation "N" if the point is in the Northern hemisphere and the designation "S" if it is in the Southern hemisphere. So, if $\angle TOE = 30°$ and T is in the Northern hemisphere, we say that its latitude is 30°N.

Note that the equator of the Earth is determined by the rotation of the Earth, whereas the prime meridian is an arbitrary choice. It was originally established by Sir George Airy as the Greenwich Meridian in 1851, and was established as the international standard by the International Meridian Conference in 1884.

It is now marked by a laser beam pointing towards the North Pole from the Royal Observatory in Greenwich, London.

Problem 5.24:

(a) What is the relationship between spherical coordinates and the latitude and longitude that we use to identify the location of points on the surface of the Earth?

(b) Why do we use a system much like spherical coordinates for points on the Earth's surface rather than using rectangular coordinates?

(c) Petr has lived his whole life in Petropavlovsk-Kamchatsky on the frigid northeastern shores of Siberia. He is sick of the cold, and has spent the last 5 years building a sailboat he hopes to sail to Honolulu, Hawaii, where he will be much warmer. If he sails at 10 km per hour, how many days will it take him to sail from Petropavlovsk-Kamchatsky, which has coordinates 52.92° N, 158.49° E, to Honolulu, which has coordinates 21.31° N, 157.83° W? You may assume the Earth is a sphere with radius 6370 kilometers, and that Petr travels along a single great circle of the sphere throughout his trip.

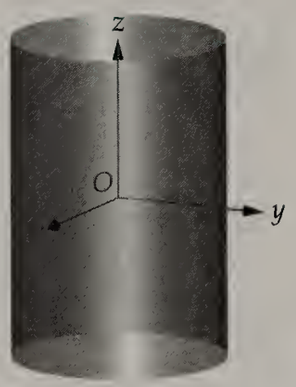
Problem 5.18:

(a) Describe the graph of all points with cylindrical coordinates (r, θ, z) such that $r = 7$. Express the equation in rectangular coordinates.

(b) Describe the graph of the points in space such that $\theta = \frac{\pi}{3}$ in cylindrical coordinates. Express the equation in rectangular coordinates.

Solution for Problem 5.18:

(a) In two dimensions, the graph of $r = 7$ is a circle centered at the origin with radius 7. In three dimensions, we can let z be any real number when we produce the graph of $r = 7$. Therefore, every cross-section of the graph perpendicular to the z-axis is a circle with radius 7 centered at a point on the z-axis. Since all these circles have the same radius, the resulting graph of $r = 7$ in three dimensions is a cylinder that extends forever in both directions of its axis. The radius of the cylinder is 7, and the z-axis is the axis of the cylinder.

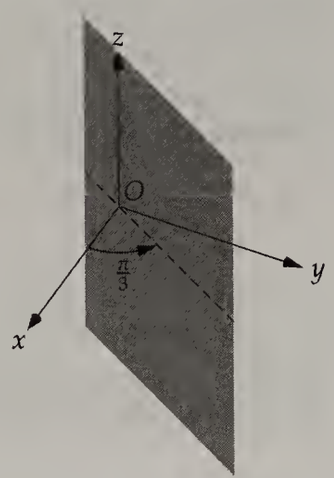


Just as in polar coordinates, we have

$$x^2 + y^2 = (r\cos\theta)^2 + (r\sin\theta)^2 = r^2(\cos^2\theta + \sin^2\theta) = r^2.$$

So, to express the equation $r = 7$ in rectangular coordinates, we square the equation to get $r^2 = 49$, which gives $x^2 + y^2 = 49$.

(b) In polar coordinates, the graph of $\theta = \frac{\pi}{3}$ is a line through the origin that makes an angle of $\frac{\pi}{3}$ with the x-axis such that the line passes through the first and third quadrants. Since we can use any value for the z-coordinate in the graph of $\theta = \frac{\pi}{3}$ in three dimensions, the graph of consists of the aforementioned line in the plane $z = 0$, as well as every point "above" or "below" (i.e., with any z-coordinate) the line in space. In other words, the graph is a plane, as shown at right.



To convert the equation to rectangular coordinates, we note that $x = r\cos\theta = r\cos\frac{\pi}{3} = \frac{r}{2}$ and $y = r\sin\theta = \frac{r\sqrt{3}}{2}$. For all $r \neq 0$, we can divide these equations to get $\frac{x}{y} = \frac{1}{\sqrt{3}}$. Rearranging this equation gives $y = \sqrt{3}x$ as the equation in rectangular form. (Note that points of the form $(0, 0, z)$ are on the graph and satisfy this equation.) □

Problem 5.19: Suppose (ρ, θ, ϕ) are the spherical coordinates of point P, where $\rho \geq 0$, and that (x, y, z) are the rectangular coordinates of point P.

(a) Find ρ in terms of x, y, and z.

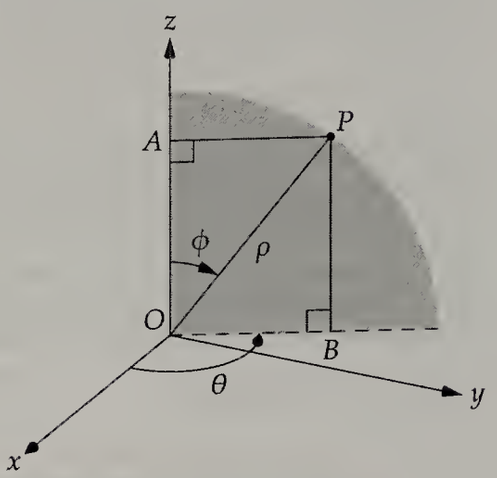
(b) Let O be the origin and A be the foot of the altitude from P to the z-axis. Find AO in terms of ρ, θ, and/or ϕ.

(c) Find x, y, and z in terms of ρ, θ, and ϕ.

Solution for Problem 5.19:

(a) The value of ρ is the distance from the origin to the point (ρ, θ, ϕ) in spherical coordinates. The distance from (x, y, z) to the origin is $\sqrt{x^2 + y^2 + z^2}$. Therefore, we must have $\rho = \sqrt{x^2 + y^2 + z^2}$.

(b) In the diagram at right, we have drawn $\overline{PA}$ perpendicular to the z-axis and $\overline{PB}$ perpendicular to the xy-plane, thereby creating rectangle $PBOA$. In right triangle POA, we have $\cos\phi = \frac{AO}{PO} = \frac{AO}{\rho}$, so $AO = \rho\cos\phi$.



(c) Since the z-coordinate of A matches that of P, we have $z = AO = \rho\cos\phi$. To find the x- and y-coordinates of P, we focus on point B in the xy-plane. Since $PBOA$ is a rectangle, we have $OB = PA$. From right triangle POA, we have $\frac{PA}{PO} = \sin\phi$, so $PA = PO\sin\phi = \rho\sin\phi$. Therefore, we have $OB = PA = \rho\sin\phi$. The x- and y-coordinates of P match those of B. The cylindrical coordinates of B are $(OB, \theta, 0)$, so the rectangular coordinates are $(OB\cos\theta, OB\sin\theta, 0)$. Therefore, the x- and y-coordinates of B and P are $x = OB\cos\theta = \rho\sin\phi\cos\theta$ and $y = OB\sin\theta = \rho\sin\phi\sin\theta$.

In summary, we have:

Important: The point with spherical coordinates (ρ, θ, ϕ) has rectangular coordinates

$$x = \rho\sin\phi\cos\theta,$$
$$y = \rho\sin\phi\sin\theta,$$
$$z = \rho\cos\phi.$$

Perhaps you've noticed that our proof only works if $0 < \phi < \frac{\pi}{2}$. You'll complete the proof for other values of ϕ as an Exercise.

□

Problem 5.20:

(a) Convert the cylindrical coordinates $\left(4, \frac{\pi}{3}, 4\right)$ to rectangular and spherical coordinates.

(b) Convert the spherical coordinates $\left(7, \frac{7\pi}{4}, \frac{\pi}{3}\right)$ to rectangular and cylindrical coordinates.

Solution for Problem 5.20:

(a) First, we convert to rectangular coordinates. We have $x = r\cos\theta = 4\cos\frac{\pi}{3} = 2$ and $y = r\sin\theta = 4\sin\frac{\pi}{3} = 2\sqrt{3}$. The z-coordinate is the same in rectangular and polar coordinates, so the rectangular coordinates of the point with cylindrical coordinates $(4, \frac{\pi}{3}, 4)$ are $(2, 2\sqrt{3}, 4)$.

Turning to spherical coordinates, the θ-coordinate is the same in both cylindrical and spherical coordinates. We can use the rectangular coordinates to find

$$\rho = \sqrt{x^2 + y^2 + z^2} = \sqrt{4 + 12 + 16} = 4\sqrt{2}.$$

Finally, since $z = \rho\cos\phi$, we have $\cos\phi = \frac{z}{\rho} = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$. So, we can let $\phi = \frac{\pi}{4}$, and $\left(4\sqrt{2}, \frac{\pi}{3}, \frac{\pi}{4}\right)$ are spherical coordinates of the point.

(b) As in part (a), we start by finding the rectangular coordinates:

$$x = \rho\sin\phi\cos\theta = 7\left(\frac{\sqrt{3}}{2}\right)\left(\frac{\sqrt{2}}{2}\right) = \frac{7\sqrt{6}}{4},$$
$$y = \rho\sin\phi\sin\theta = 7\left(\frac{\sqrt{3}}{2}\right)\left(-\frac{\sqrt{2}}{2}\right) = -\frac{7\sqrt{6}}{4},$$
$$z = \rho\cos\phi = 7\left(\frac{1}{2}\right) = \frac{7}{2}.$$

To convert these to cylindrical coordinates, we note that we already have the z-coordinate from the rectangular coordinates and the θ-coordinate from the spherical coordinates. Finally, we have

$$r = \sqrt{x^2 + y^2} = \sqrt{\left(\frac{7\sqrt{6}}{4}\right)^2 + \left(\frac{7\sqrt{6}}{4}\right)^2} = \frac{7\sqrt{6}}{4}\cdot\sqrt{2} = \frac{7\sqrt{3}}{2},$$

so $\left(\frac{7\sqrt{3}}{2}, \frac{7\pi}{4}, \frac{7}{2}\right)$ are cylindrical coordinates of the point.

□

Problem 5.21: Convert each of the following rectangular coordinates to cylindrical coordinates and to spherical coordinates.

(a) $(6, -6, -2\sqrt{6})$ (b) $(-4, 0, 4\sqrt{3})$

Solution for Problem 5.21:

(a) In two dimensions, the point with rectangular coordinates $(6, -6)$ has polar coordinates $\left(6\sqrt{2}, \frac{7\pi}{4}\right)$, so the point in space with rectangular coordinates $(6, -6, -2\sqrt{6})$ has cylindrical coordinates $\left(6\sqrt{2}, \frac{7\pi}{4}, -2\sqrt{6}\right)$.

Turning to spherical coordinates, we have $\rho = \sqrt{x^2 + y^2 + z^2} = \sqrt{36 + 36 + 24} = 4\sqrt{6}$. The θ-coordinate is the same in spherical coordinates as in cylindrical coordinates. To find ϕ, we use the fact that $z = \rho\cos\phi$, so we have $-2\sqrt{6} = (4\sqrt{6})\cos\phi$. Therefore, we have $\cos\phi = -\frac{1}{2}$, so we take $\phi = \frac{2\pi}{3}$. This gives us the spherical coordinates $\left(4\sqrt{6}, \frac{7\pi}{4}, \frac{2\pi}{3}\right)$.

(b) In two dimensions, the point with rectangular coordinates $(-4, 0)$ has polar coordinates $(4, \pi)$, so the the point in three dimensions with rectangular coordinates $(-4, 0, 4\sqrt{3})$ has cylindrical coordinates $(4, \pi, 4\sqrt{3})$.

In spherical coordinates, we have $\rho = \sqrt{x^2 + y^2 + z^2} = \sqrt{16 + 48} = 8$, and the θ-coordinate is the same in spherical coordinates as in cylindrical coordinates. From $z = \rho\cos\phi$, we have $4\sqrt{3} = 8\cos\phi$, so $\cos\phi = \frac{\sqrt{3}}{2}$ and $\phi = \frac{\pi}{6}$. Therefore, $\left(8, \pi, \frac{\pi}{6}\right)$ are spherical coordinates of the point.

□

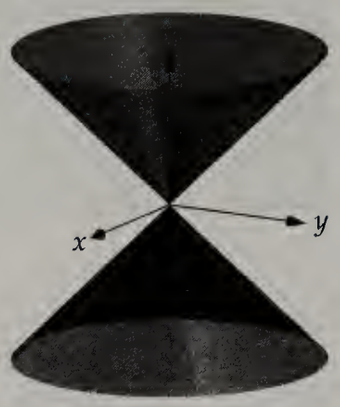
Problem 5.22:

(a) Describe the graph of the equation $z^2 = x^2 + y^2$.

(b) Find parametric equations for x, y, and z for the equation $z^2 = x^2 + y^2$. (You will need *two* parameters.)

Solution for Problem 5.22:

(a) We can begin by experimenting with different values of z. When $z = 0$, we have $x^2 + y^2 = 0$, so when $z = 0$, the graph is just a point. Letting $z = 1$ or $z = -1$ gives $x^2 + y^2 = 1$, which describes a circle with radius 1. Similarly, for larger and larger values of z, the resulting equations have graphs that are circles with larger and larger radii.



We might therefore guess that the graph is a pair of cones that open forever in both directions. Expressing the surface in cylindrical coordinates makes it very obvious that this is the case. We have $z^2 = x^2 + y^2 = r^2$, so $z = \pm r$. Therefore, the radius of each circular cross-section of the graph equals the distance from the cross-section to the origin. The graph is the pair of cones shown at right. Each cone has the z-axis as its axis.

(b) Consider the upper cone, for which $z > 0$. In part (a), we saw that every cross-section of the surface with constant z-coordinate is a circle with radius r centered at a point on the z-axis, where $r = z$. So, the x- and y-coordinates can be parameterized as $x = r\cos\theta$, $y = r\sin\theta$. This holds for any positive value of z, so we can parameterize the upper cone with $x = r\cos\theta$, $y = r\sin\theta$, $z = r$.

This parameterization holds for the lower cone, too! If we have $x = r\cos\theta$, $y = r\sin\theta$, and $z = r$ then $x^2 + y^2 = r^2 = z^2$ still, so $(x, y, z) = (r\cos\theta, r\sin\theta, r)$ is on the circle centered at $(0, 0, r)$ in the plane $z = r$. As θ ranges from 0 to 2π, we trace out the whole circle, so the parameterization $x = r\cos\theta$, $y = r\sin\theta$, $z = r$ generates all the points on the lower cone for $r < 0$. Finally, taking $r = 0$ shows that the parameterization includes the origin, as well.

□

One significant difference in our answer to part (b) above and the parameterizations we performed in Section 5.1 is that our parameterization in part (b) above required 2 parameters, whereas each parameterization in Section 5.1 required only 1 parameter. The reason for this difference is that our parameterizations in Section 5.1 were used to describe *1-dimensional curves in a 2-dimensional plane*, whereas the parameterization of the cone describes a *2-dimensional surface in a 3-dimensional space*. Since we are describing a 2-dimensional surface, we need two parameters.

Problem 5.23:

(a) Describe the graph of all points with spherical coordinates (ρ, θ, ϕ) such that $\rho = 5$. Express the equation in rectangular coordinates, and give a parameterization of the equation for x, y, and z.

(b) Describe the graph of the points in space such that $\phi = \frac{2\pi}{3}$ in spherical coordinates. Express the equation in rectangular coordinates.

Solution for Problem 5.23:

(a) The graph of $\rho = 5$ consists of all points that are exactly 5 units from the origin. Therefore, the graph is a sphere centered at the origin with radius 5. To express the equation in rectangular coordinates, we write ρ as $\sqrt{x^2 + y^2 + z^2}$. Since this quantity is always nonnegative, we can square the equation $\sqrt{x^2 + y^2 + z^2} = 5$ to produce the more attractive equivalent equation $x^2 + y^2 + z^2 = 25$.

We can parameterize this equation using the conversion from spherical to rectangular coordinates. Since $\rho = 5$, we have the parameterization $x = \rho\cos\theta\sin\phi = 5\cos\theta\sin\phi$, $y = \rho\sin\theta\sin\phi = 5\sin\theta\sin\phi$, $z = \rho\cos\phi = 5\cos\phi$.

(b) Let O be the origin and P be a point that has $\phi = \frac{2\pi}{3}$ in spherical coordinates. The ϕ-coordinate gives us the angle that $\overline{OP}$ makes with the positive portion of the z-axis. Therefore, any point on the ray $\overrightarrow{OP}$ has $\phi = \frac{2\pi}{3}$ when expressed in spherical coordinates. Moreover, we can rotate any point on this ray about the z-axis

by any angle to produce another point with the same ϕ-coordinate. In other words, we can form a portion of the graph by rotating the whole ray around the z-axis, thereby sweeping out a cone. Since we can allow ρ to be negative, the graph consists of two cones, one below the xy-plane (for $\rho > 0$), and one above the xy-plane (for $\rho < 0$).

We can see for certain that the graph is a cone by considering a particular cross-section perpendicular to the z-axis. For such a cross-section, z is fixed. Because ϕ is also fixed by the equation $\phi = \frac{2\pi}{3}$, the relationship $z = \rho\cos\phi$ tells us that ρ must also be constant for all points on the graph in this cross-section. Moreover, since ϕ is fixed for all points in the graph, the relationship $z = \rho\cos\phi$ tells us that ρ is directly proportional to z. Finally, for any point in the cross-section, we have

$$x^2 + y^2 = (\rho\sin\phi\cos\theta)^2 + (\rho\sin\phi\sin\theta)^2 = (\rho^2\sin^2\phi)(\cos^2\theta + \sin^2\theta) = \rho^2\sin^2\phi.$$

So, the cross-section is a circle with radius $\rho\sin\phi$, which is directly proportional to ρ, and therefore directly proportional to z. Since every cross-section is a circle with radius directly proportional to the magnitude of z, the graph of $\phi = \frac{2\pi}{3}$ is a pair of cones with vertex at the origin and with the z-axis as their common axis.

We can express the equation in rectangular coordinates by starting from the equation $x^2 + y^2 = \rho^2\sin^2\phi$ above. Since $\phi = \frac{2\pi}{3}$, we have $x^2 + y^2 = \frac{3\rho^2}{4}$. We also have $\rho^2 = x^2 + y^2 + z^2$, so $x^2 + y^2 = \frac{3}{4}(x^2 + y^2 + z^2)$, which simplifies as $x^2 + y^2 = 3z^2$. (We also could have noted that $z = \rho\cos\phi$, so $\rho\sin\phi = z\frac{\sin\phi}{\cos\phi} = z\tan\phi$. For $\phi = \frac{2\pi}{3}$, we then have $x^2 + y^2 = \rho^2\sin^2\phi = z^2\tan^2\phi = 3z^2$.)

□

Problem 5.24:

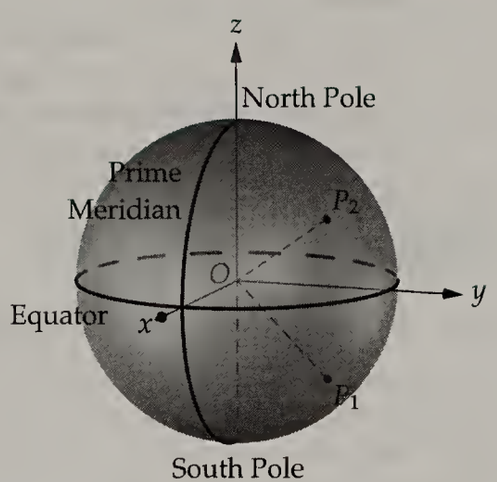
(a) What is the relationship between spherical coordinates and the latitude and longitude that we use to identify the location of points on the surface of the Earth? (A definition of latitude and longitude can be found on page 180.)

(b) Why do we use a system much like spherical coordinates for points on the Earth's surface rather than using rectangular coordinates?

(c) Petr has lived his whole life in Petropavlovsk-Kamchatsky on the frigid northeastern shores of Siberia. He is sick of the cold, and has spent the last 5 years building a sailboat he hopes to sail to Honolulu, Hawaii, where he will be much warmer. If he sails at 10 km per hour, how many days will it take him to sail from Petropavlovsk-Kamchatsky, which has coordinates 52.92° N, 158.49° E, to Honolulu, which has coordinates 21.31° N, 157.83° W? You may assume the Earth is a sphere with radius 6370 kilometers, and that Petr travels along a single great circle of the sphere throughout his trip.

Solution for Problem 5.24:

(a) Let the center of the Earth be the origin, and the Earth's axis be the z-axis, with the North Pole on the positive z-axis and the South Pole on the negative z-axis. Then, longitude is closely related to θ and latitude is related to ϕ. If we take the prime meridian to be entirely within in the xz-plane, so that the prime meridian intersects the positive x-axis as shown, then we can view "east" as counterclockwise, so the longitude of a point in the Eastern hemisphere equals the angle θ in spherical coordinates. The points in the Western hemisphere correspond to those with values of θ for which $180° < \theta < 360°$. We can take care of these by noting that subtracting 360° from θ doesn't change the location of a point with this θ-coordinate, and $-180° < \theta - 360° < 0°$. Therefore, the θ-coordinate of any point in the Western hemisphere equals the negative of its longitude.



Turning to latitude, suppose P_1 is a point on the Earth's surface in the Southern hemisphere and O is the origin. The angle between $\overline{OP_1}$ and the positive z-axis (ϕ in spherical coordinates) is 90° plus the latitude of P_1. Meanwhile, suppose point P_2 is on the Earth's surface in the Northern hemisphere. The angle between $\overline{OP_2}$ and the positive z-axis is 90° minus the latitude of P_2.

(b) If we center the Earth at the origin (and assume the Earth is a sphere), then in spherical coordinates, all the points on the Earth's surface have the same ρ coordinate. So, in order to tell anyone the location of a point on the Earth's surface, we need only tell them two coordinates, θ and ϕ. In rectangular coordinates, all three coordinates must be specified to determine a point on the Earth's surface.

(c) Let the center of the Earth be O, let Petropavlovsk-Kamchatsky be P, and let Honolulu be H. To find the amount of time Petr spends sailing, we must find the distance along the surface of Earth from P to H. This distance is the arc of a great circle. We know the radius of this circle; it equals the radius of the Earth. We only have to find the angle swept out by this arc, which is $\angle POH$, and then we can find the length of the arc. Unfortunately, it's not so obvious how to find that angle from the latitude and longitude of P and of H.

However, we do know that PO and OH equal the radius of the Earth, which we'll let be r. If we can determine any other angle or the third side length of $\triangle POH$, then we can find $\angle POH$. We are essentially given the spherical coordinates of P and of H; we can use these to find the distance from P to H using the three-dimensional distance formula.

We let $(\rho_P, \theta_P, \phi_P)$ and $(\rho_H, \theta_H, \phi_H)$ be the spherical coordinates of P and H. The first coordinate is easy; it equals the radius of the Earth, so we have $\rho_P = \rho_H = r$. Turning to the θ-coordinate, we apply the reasoning of part (a) to the longitudes of P and H. Since P is in the Eastern hemisphere, its θ-coordinate equals its longitude, so $\theta_P = 158.49°$ for P. Point H is in the Western hemisphere, so we can take its θ to be the negative of its longitude, which gives $\theta_H = -157.83°$. Finally, both P and H are in the Northern hemisphere, so their ϕ-coordinates are in 90° minus their latitudes. This gives us $\phi_P = 37.08°$ and $\phi_H = 68.69°$. Now, we can find PH by converting these spherical coordinates to rectangular coordinates and using the distance formula. We can do both of these in one step using the relationships $x = \rho\cos\theta\sin\phi$, $y = \rho\sin\theta\sin\phi$, and $z = \rho\cos\phi$. Since $\rho_P = \rho_H = r$, we have

$$PH = \sqrt{(r\cos\theta_H\sin\phi_H - r\cos\theta_P\sin\phi_P)^2 + (r\sin\theta_H\sin\phi_H - r\sin\theta_P\sin\phi_P)^2 + (r\cos\phi_H - r\cos\phi_P)^2}.$$

Substituting all our values from above, we find that $PH \approx 4970$ km. Since we have PO, HO, and PH, we can use the Law of Cosines to find $\cos\angle POH$:

$$\cos\angle POH = \frac{PO^2 + HO^2 - PH^2}{2(PO)(OH)} \approx 0.696.$$

Therefore, $\angle POH \approx 45.9°$, which means that Petr will sail $\frac{45.9°}{360°} \approx 0.128$ of a great circle during his trip. Since the circumference of the Earth is $2\pi(6370) \approx 40000$ km, Petr will sail $0.128\,(40000) \approx 5120$ km. He sails at 10 kilometers per hour, so he'll sail for 512 hours, which is 21 days and 8 hours.

That's a long trip. You might wonder if there's a formula to find the distance between two points on the Earth given their latitudes and longitudes. There is, and you'll prove it on page 480.

□

Exercises

5.3.1 Express each of the following rectangular coordinates in cylindrical and spherical coordinates.

(a) $(-2\sqrt{2}, 2\sqrt{2}, 4)$ (b) $(3, 0, -3\sqrt{3})$

5.3.2 Suppose the point with spherical coordinates (ρ, θ, ϕ), where $\rho \geq 0$, has cylindrical coordinates (r, θ, z). Express ρ in terms of r, θ, and z.

5.3.3 In Problem 5.19 we only proved the relationship between spherical coordinates and rectangular coordinates for $0 < \phi < \frac{\pi}{2}$. Complete our proof by addressing points for which $\frac{\pi}{2} \leq \phi \leq \pi$ or $\phi = 0$.

5.3.4 Show that if $\rho \geq 0$, then the distance from $(\rho \sin\phi \cos\theta, \rho \sin\phi \sin\theta, \rho \cos\phi)$ to the origin is ρ.

5.3.5 Let the cylindrical coordinates of P be (r, θ, z). What are cylindrical coordinates of the following points:

(a) The reflection of P through the xy-plane.

(b) The rotation of P by π around the z-axis.

(c)★ The reflection of P through the origin. (The reflection of a point X through a point Y is the point Z such that Y is the midpoint of $\overline{XZ}$.)

5.3.6 We showed that the graph of $\phi = \frac{2\pi}{3}$ in spherical coordinates is a pair of cones. Are there any values of c for which the graph of $\phi = c$ is *not* a pair of cones?

5.3.7 Suppose we used cylindrical coordinates to identify the location of points on the surface of the Earth, with the center of the Earth as the origin.

(a) Would it be possible to travel along the surface of the Earth and keep the r-coordinate constant? The θ-coordinate? The z-coordinate?

(b) Is it possible to travel along the surface of the Earth and keep *two* of the cylindrical coordinates constant?

5.3.8

(a) Describe the graph of all points such that $\rho = 4$ and $\theta = \frac{2\pi}{5}$ in spherical coordinates.

(b)★ Describe the graph of all points such that $\theta = \frac{\pi}{2}$ and $\rho = 6\sin\phi$ in spherical coordinates.

5.3.9 Consider the equation $(x+3)^2 + (y+5)^2 + (z-1)^2 = 4$.

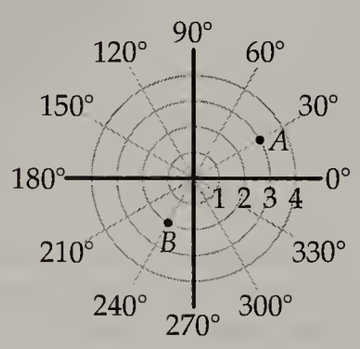
(a) Find a parameterization for x, y, and z for all points that satisfy the equation.

(b)★ Find a parameterization for the portion of the graph of the equation that is on or above the xy-plane.

5.4 Summary

Rather than having a single equation that relates x and y, we sometimes define a relationship between x and y through an intermediary variable called a **parameter**. We do so by expressing x and y as functions of a third variable, such as t. We call the system $x = f(t)$, $y = g(t)$ a pair of **parametric equations**, where t is the parameter. When we graph parametric equations $x = f(t)$, $y = g(t)$, we graph all points (x, y) such that there is a value of t for which $x = f(t)$ and $y = g(t)$.

In addition to the more common **rectangular coordinates** (x, y), **polar coordinates** offer another method for denoting the location of a point on a plane. As with rectangular coordinates, we start with a point we call the origin and we identify the location of each point with an ordered pair, (r, θ). The r-coordinate is the distance from the point to the origin; we call this coordinate the **radius**. The θ-coordinate is the **angle**, which we define in the same way we related angles to points on the unit circle.



We can also use a negative radius, which means "go in the direction opposite the one indicated by θ." For example, we could express A as $(-3, 210°)$, which tells us that A is 3 units in the direction opposite 210°, which is the direction of 30°. Similarly, the polar coordinates (r, θ) and $(-r, \theta + 180°)$ always refer to the same point. We can also use a negative angle, so point B at right could be expressed as $(2, -120°)$.

If (x, y) in rectangular coordinates is the same point as (r, θ) in polar coordinates, then

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad r^2 = x^2 + y^2, \qquad \tan\theta = \frac{y}{x},$$

where $\tan\theta = \frac{y}{x}$ holds only if $x \neq 0$. If $x = 0$ and $y \neq 0$, then $\cos\theta = 0$ and $r^2 = y^2$. If $x = y = 0$, then $r = 0$ and θ can be anything.

Just as we use the horizontal and vertical coordinates (x, y) to identify the location of a point in the two-dimensional Cartesian plane, we can identify the location of a point in three-dimensional space with three coordinates (x, y, z). We call the plane in which $z = 0$ the xy-plane. Points on this plane are identified by the coordinates $(x, y, 0)$, where the x and y have the horizontal and vertical meanings they have in the Cartesian plane. For other points (x, y, z) in space, the z-coordinate indicates how far above $(z > 0)$ or below $(z < 0)$ the xy-plane the point is. The x- and y-coordinates of the point are the same as those of the point in the $z = 0$ plane that is directly above or below the point (x, y, z).

Important: The distance between the points (x_1, y_1, z_1) and (x_2, y_2, z_2) is

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

There are two commonly used coordinate systems in three dimensions that use angles among their coordinates. In **cylindrical coordinates**, the x- and y-coordinates are replaced with r- and θ-coordinates such that the point in the Cartesian plane with rectangular coordinates (x, y) has polar coordinates (r, θ). The third cylindrical coordinate is the same as the z-coordinate from rectangular coordinates. A depiction of the point (r, θ, z) in cylindrical coordinates is shown at left below. When expressing a point in polar coordinates, the radius always comes first, the angle second, and the z-coordinate last.

If the cylindrical coordinates (r, θ, z) represent the same point as the rectangular coordinates (x, y, z), then $x = r\cos\theta$ and $y = r\sin\theta$. As with polar coordinates, we usually choose r to be nonzero and θ such that $0 \leq \theta < 2\pi$. However, r can be negative and θ can be outside the interval $[0, 2\pi)$, again just as with polar coordinates.

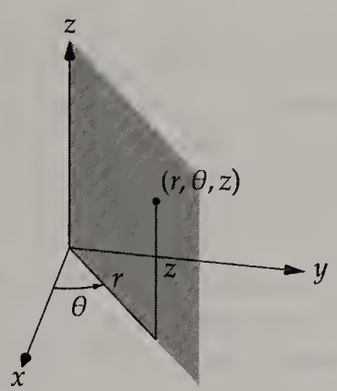


Figure 5.3: Cylindrical Coordinates

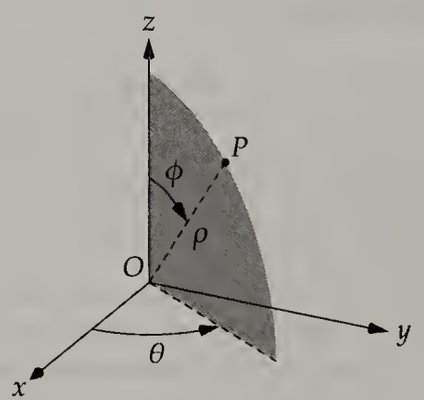


Figure 5.4: Spherical Coordinates

The **spherical coordinates** of a point are typically denoted as (ρ, θ, ϕ), The ρ-coordinate is the distance from the P to the origin. The θ-coordinate is essentially the same as the angle in cylindrical coordinates. The ϕ-coordinate is the angle between $\overline{OP}$ and the positive z-axis, as shown. We typically choose the spherical coordinates such that $\rho \geq 0$, $0 \leq \theta \leq 2\pi$, and $0 \leq \phi \leq \pi$. The point with spherical coordinates (ρ, θ, ϕ) has rectangular coordinates

$$x = \rho\sin\phi\cos\theta, \qquad y = \rho\sin\phi\sin\theta, \qquad z = \rho\cos\phi,$$

and we have $\rho^2 = x^2 + y^2 + z^2$.

Things To Watch Out For!

WARNING!!

- If parametric equations $x = f(t)$, $y = g(t)$ satisfy an equation in terms of x and y, we only know that the graph of the parametric equations is *part of* the graph of the non-parametric equation. To show that the graph of the parametric equations is the same as the graph of the non-parametric equation, we must show that every point in the graph of the non-parametric equation is in the graph of the parametric equations. We can usually do so by showing that for each point (x_1, y_1) on the graph of the non-parametric equation, there is a value of t such that $x_1 = f(t)$, $y_1 = g(t)$.
- The graph of parametric equations is not necessarily the same as that of a non-parametric equation satisfied by the parametric equations. Sometimes the parametric equations have restrictions that limit the graph of the parametric equations to only a portion of the graph of the non-parametric equation.
- When reading about spherical coordinates in another source, make sure you know what each variable stands for and what the restrictions on each are. While most sources will define the two angles in spherical coordinates in essentially the same way we have here, some will use different variables than we have. Even more confusing, some will use the same variables, but reverse their meanings. Furthermore, some will restrict ρ to be nonnegative, and some will use r rather than ρ.

REVIEW PROBLEMS

5.25 Find the graph of the parametric equations $x = 3a - 1$, $y = 2a^2 - a$.

5.26 Find parametric equations for x and y such that $x^2 + 4x + y^2 - 2y = 11$.

5.27 Find the graph of the parametric equations $x = (\sin t)(\cos t)$, $y = 2\cos^2 2t$.

5.28 Describe the curve that results from graphing the parametric equations $x = \frac{1}{1-t^2}$, $y = \frac{t}{1-t^2}$.

5.29 Convert each of the following rectangular coordinates to polar coordinates. You may use a calculator on this problem. Approximate angles to the nearest hundredth of a radian.

(a) $(-7, -7)$ (b) $(6\sqrt{3}, -6\sqrt{2})$ (c) $(-5, 8)$

5.30 Convert each of the following polar coordinates to rectangular coordinates.

(a) $(5, 0°)$ (b) $(-4, 210°)$ (c) $\left(8, -\frac{\pi}{4}\right)$

5.31 Describe the graph of the equation $r^2 = 3r\cos\theta - 5r\sin\theta - 2$.

5.32

(a) What point results from rotating $(4, 5)$ by 60° clockwise about the origin?

(b) What point results from rotating $(4, 5)$ by 60° clockwise about $(2, 5)$?

(c)★ What point results from rotating $(4, 5)$ by 60° clockwise about $(-2, 7)$?

5.33 Let point M be the midpoint of the segment with endpoints (r_1, θ_1) and (r_2, θ_2) in polar coordinates.

(a) Must it always be possible to express M in polar coordinates as $\left(\frac{r_1+r_2}{2}, \theta\right)$ for some angle θ?

(b) Must it always be possible to express M in polar coordinates as $\left(r, \frac{\theta_1+\theta_2}{2}\right)$ for some radius r?

5.34 What symmetry must the graph of $f(r, \theta) = 0$ have if:

(a) $f(r, \theta) = f(-r, \theta)$ for all r and θ.

(b) $f(r, \theta) = f(r, -\theta)$ for all r and θ.

5.35 Express each of the following rectangular coordinates in cylindrical and spherical coordinates.

(a) $(2, -2, 2\sqrt{2})$ (b) $(\sqrt{3}, 3, -2\sqrt{3})$

5.36

(a) Convert the cylindrical coordinates $\left(6, \frac{2\pi}{3}, -1\right)$ to rectangular coordinates.

(b) Convert the spherical coordinates $\left(12, \frac{\pi}{4}, \frac{3\pi}{4}\right)$ to rectangular coordinates.

5.37 Describe the graph of all points in space such that $y^2 + z^2 = 9$.

5.38 What is the graph of $r + z = 6$ in cylindrical coordinates?

5.39 Describe the graphs of each of the following equations:

(a) $r = 5$ in cylindrical coordinates.

(b) $\rho = 3$ in spherical coordinates.

(c) $\theta = -\frac{\pi}{3}$ in cylindrical coordinates.

(d) $\phi = \frac{3\pi}{4}$ in spherical coordinates.

5.40 Describe the surface that is the graph of all points with cylindrical coordinates (r, θ, z) such that $r^2 + z^2 = 2z$.

5.41 Describe the surface that consists of all points with spherical coordinates (ρ, θ, ϕ) such that $\rho = 6\cos\phi$.

5.42 Describe the graph of $\rho = \frac{3}{\sin\phi}$ in spherical coordinates.

5.43

(a) What geometric transformation maps any point with spherical coordinates (ρ, θ, ϕ) to a point with spherical coordinates $(\rho, \theta + \pi, \phi)$?

(b) What geometric transformation maps any point with spherical coordinates (ρ, θ, ϕ) to a point with spherical coordinates $(\rho, \theta, \pi - \phi)$?

Challenge Problems

5.44

(a) Graph $r = \sin(2\theta)$.

(b) Graph $r = \sin(3\theta)$.

(c) Graph $r = \sin(4\theta)$.

(d) Graph $r = \sin(5\theta)$.

(e)★ Generalize your results from the first four parts: describe the graph of $r = \sin(k\theta)$ for any integer k.

5.45

(a) Find parametric equations for x and y whose graph is a circle with center (h, k) and radius r.

(b) Find parametric equations for x, y, and z whose graph is a sphere with center (a, b, c) and radius ρ. (You will need two parameters.)

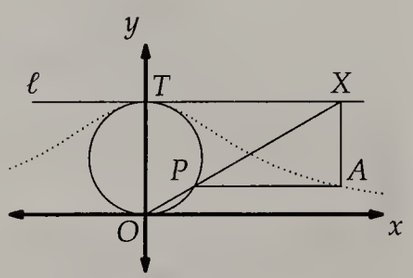
(c) Find parametric equations for x, y, and z whose graph is a cylinder with radius r and with the x-axis as its axis. (You will need two parameters.)

5.46 Find parametric equations whose graph is the portion of the curve $x^2 - 4x + y^2 - 2y + 1 = 0$ that is above the x-axis. **Hints:** 177, 107

5.47 Let (x, y) be a point that does not lie on the negative x-axis, and let (r, θ) be the point's polar coordinates, where $r > 0$ and $-\pi < \theta < \pi$. Show that $\theta = 2 \arctan \frac{y}{x + \sqrt{x^2 + y^2}}$. **Hints:** 66

5.48 Find the smallest distance between the point $(1, 2, 3)$ and a point on the graph of the parametric equations $x = 2 - t$, $y = 4 + t$, $z = 3 + 2t$.

5.49 The **witch of Agnesi**, named after mathematician Maria Gaetana Agnesi, is a curve that is defined as follows. Let O be the origin, let T be the point $(0, 1)$, and let ℓ be the line through T parallel to the x-axis. Let C be the circle centered at $\left(0, \frac{1}{2}\right)$ with radius $\frac{1}{2}$. For any point P on the circle besides O and T, we draw a ray from O through P. Let this ray intersect ℓ at point X. We then draw the altitude from X to the line through P parallel to the x-axis. The foot of this altitude, point A, is on the witch curve. When we trace out the resulting points A for all possible P, and include point T, we get the witch curve, which is dotted in the diagram.



(a) Find a parameterization for the witch curve. **Hints:** 23

(b) Find a function f such that the witch curve is the graph of the function $y = f(x)$.

5.50★ A sphere with radius 2 centered at the origin is painted red. A hole with radius 1 is drilled through the sphere along the x-axis. Find a parameterization for the red surface that remains. **Hints:** 244, 10, 235

5.51★ Coin $\mathcal{A}$ with radius 1 is placed on the Cartesian plane so that its center is at the origin. Coin $\mathcal{B}$ with radius 1 is placed on the Cartesian plane so that its center is at $(0, 2)$, which means that the two coins are tangent at $(0, 1)$. Suppose the point on $\mathcal{B}$ that is initially touching coin $\mathcal{A}$ is painted red. We then roll coin $\mathcal{B}$ counterclockwise around coin $\mathcal{A}$ such that the two always remain touching, and coin $\mathcal{A}$ never moves.

(a) Find parametric equations for the path of the red point. **Hints:** 148, 117, 8

(b) During a full revolution of coin $\mathcal{B}$ around coin $\mathcal{A}$, how many revolutions does the red point make around the center of coin $\mathcal{B}$? (The point completes a full revolution about the center of coin $\mathcal{B}$ whenever it is again directly below the center of $\mathcal{B}$.)

(c) Suppose the radius of coin $\mathcal{A}$ is n, where n is a positive integer, and that coin $\mathcal{B}$ has radius 1 and is initially centered at $(0, n + 1)$. Answer the questions of part (a) and part (b) if we again roll coin $\mathcal{B}$ about the circumference of coin $\mathcal{A}$.

There can be very little of present-day science and technology that is not dependent on complex numbers in one way or another. – Keith Devlin

CHAPTER 6

Basics of Complex Numbers

The numbers we have used so far in this book are **real numbers**. These are the numbers that we can plot on the number line. One characteristic of a real number is that its square is nonnegative. In particular, there is no real number whose square is -1.

But imagine that such a number exists: that there is some number whose square is -1. This leap brings us to a rich area of mathematics that we study in the next three chapters. We use the symbol i to represent a number whose square is -1:

$$i^2 = -1.$$

A **complex number** is a number of the form $a + bi$, where a and b are real numbers and $i^2 = -1$. We call a the **real part** of $a + bi$ and we call b the **imaginary part**. If $b \neq 0$, then $a + bi$ is **nonreal**. If $a = 0$, then the complex number $a + bi$ is called an **imaginary number**. (Some sources refer to imaginary numbers as "pure imaginary.") The number 0 is both real and imaginary, and all real numbers and all imaginary numbers are complex numbers. Just as we use the symbol $\mathbb{R}$ to refer to "all real numbers," we use the symbol $\mathbb{C}$ to refer to "all complex numbers."

6.1 Complex Number Arithmetic

In this section, we cover the basic arithmetic of complex numbers. To add two complex numbers, we add their real parts and add their imaginary parts. So, the sum of the complex numbers $a + bi$ and $c + di$ as $(a + c) + (b + d)i$. We define the product of two complex numbers such that multiplication is distributive. So, we have

$$\begin{aligned}(a + bi)(c + di) &= a(c + di) + bi(c + di)\\ &= ac + adi + bci + bdi^2\\ &= ac + adi + bci - bd\\ &= (ac - bd) + (ad + bc)i.\end{aligned}$$

Problems

Problem 6.1: Find all values of z such that $z^2 + 81 = 0$.

Problem 6.2: Evaluate i^{2009}.

Problem 6.3: Let $w = 3 - 2i$ and $z = 2 + 5i$. Express each of the following as a single complex number:

(a) $w + z$

(b) $w - 3z$

(c) wz

(d) $2w^2z - 3wz^2$

Problem 6.4: Show that complex numbers are commutative under addition and multiplication. In other words, explain why $w + z = z + w$ and $wz = zw$ for all complex numbers w and z.

Problem 6.5:

(a) Find all real values of k such that the product $(3 + 2i)(3 + ki)$ is a real number.

(b) Let $z = a + bi$, where a and b are real numbers, and let $\overline{z} = a - bi$. Prove that $z \cdot \overline{z}$ is real.

Problem 6.6: Write $\dfrac{2-i}{3+2i}$ as a complex number. In other words, find real numbers a and b such that $\dfrac{2-i}{3+2i} = a+bi$.

Problem 6.7: The **multiplicative inverse** of a number z is the number w such that $wz = 1$. For example, 0.25 is the multiplicative inverse of 4. Show that any nonzero complex number $a + bi$ has a multiplicative inverse, and express that inverse in terms of a and b.

Problem 6.8: Find all complex numbers z such that $\dfrac{3z - 5i}{2z + 3} = 3 - 2i$.

Problem 6.1: Find all values of z such that $z^2 + 81 = 0$.

Solution for Problem 6.1: Isolating z gives $z^2 = -81$. We might immediately see that our solutions are $z = \pm 9i$, because $i^2 = -1$ and both 9^2 and $(-9)^2$ equal 81. We could also "take the square root" of both sides of $z^2 = -81$ and write $z = \pm\sqrt{-81}$, where $\sqrt{-81}$ is a shorthand for $i\sqrt{81}$. So, we have $z = \pm i\sqrt{81} = \pm 9i$. Another way to think about this latter approach is to think of -81 as $9^2 i^2$, since $i^2 = -1$. Therefore, we have $z^2 = (9i)^2$, from which we have $z = \pm 9i$. □

WARNING!! We have to be careful about using notation like $\sqrt{-81}$. The real-valued function $f(x) = \sqrt{x}$ is only defined for $x \geq 0$. When we write $\sqrt{-81}$ as an intermediate step to solving an equation like $z^2 = -81$, we should keep in mind that we mean $i\sqrt{81}$.

Problem 6.2: Evaluate i^{2009}.

Solution for Problem 6.2: We start by experimenting with smaller powers of i, hoping to find a pattern. We have

$i^1 = i$, $i^2 = -1$, $i^3 = i^2 \cdot i = -1 \cdot i = -i$, and $i^4 = i^2 \cdot i^2 = (-1)^2 = 1$. Then, the powers repeat:

$$\begin{aligned} i^5 &= i^4 \cdot i = i, \\ i^6 &= i^4 \cdot i^2 = i^2, \\ i^7 &= i^4 \cdot i^3 = i^3, \\ i^8 &= i^4 \cdot i^4 = i^4 = 1. \end{aligned}$$

Therefore, we see that:

Important: The powers of i repeat in cycles of 4:

$$\begin{aligned} i^1 &= i^5 = i^9 = \cdots = i, \\ i^2 &= i^6 = i^{10} = \cdots = -1, \\ i^3 &= i^7 = i^{11} = \cdots = -i, \\ i^4 &= i^8 = i^{12} = \cdots = 1. \end{aligned}$$

Because 2009 leaves a remainder of 1 when divided by 4, we have $i^{2009} = i^{4 \cdot 502 + 1} = (i^4)^{502} \cdot i^1 = i$. □

Problem 6.3: Let $w = 3 - 2i$ and $z = 2 + 5i$. Express each of the following as a single complex number:

(a) $w + z$

(b) $w - 3z$

(c) wz

(d) $2w^2z - 3wz^2$

Solution for Problem 6.3:

(a) To add complex numbers, we add the real parts and add the imaginary parts:

$$w + z = (3 - 2i) + (2 + 5i) = (3 + 2) + (-2i + 5i) = 5 + 3i.$$

(b) We have

$$w - 3z = 3 - 2i - 3(2 + 5i) = 3 - 2i - 6 - 15i = -3 - 17i.$$

(c) We multiply complex numbers using the distributive property, just as we do when we multiply binomials:

$$(3 - 2i)(2 + 5i) = 3(2 + 5i) - 2i(2 + 5i) = 6 + 15i - 4i - 10i^2 = 6 + 11i - 10(-1) = 16 + 11i.$$

(d) We save ourselves a little bit of work by factoring the expression first: $2w^2z - 3wz^2 = wz(2w - 3z)$. In the previous part, we found $wz = 16 + 11i$. We also have

$$2w - 3z = 2(3 - 2i) - 3(2 + 5i) = 6 - 4i - 6 - 15i = -19i,$$

so

$$wz(2w - 3z) = (16 + 11i)(-19i) = -304i - 209i^2 = 209 - 304i.$$

□

Sidenote: We frequently use w and z as variables to represent complex numbers, rather than the x and y that are typically used in problems about real numbers. That said, the variable x is so common that you'll see many equations in terms of x that have nonreal solutions.

Problem 6.4: Show that complex numbers are commutative under addition and multiplication. In other words, explain why $w + z = z + w$ and $wz = zw$ for all complex numbers w and z.

Solution for Problem 6.4: We let $w = a + bi$ and $z = c + di$, and we have

$$w + z = (a + bi) + (c + di) = (a + c) + (b + d)i = (c + a) + (d + b)i = z + w.$$

We also have

$$wz = (a + bi)(c + di) = (ac - bd) + (ad + bc)i,$$
$$zw = (c + di)(a + bi) = (ca - db) + (cb + da)i = (ac - bd) + (ad + bc)i,$$

so $wz = zw$. □

We can similarly show that complex numbers are associative under addition and multiplication, and that multiplication of complex numbers is distributive.

Important: For any complex numbers v, w, and z, we have:

- $w + z = z + w$ and $wz = zw$
- $v + (w + z) = (v + w) + z$ and $(vw)z = v(wz)$
- $v(w + z) = vw + vz$

There's not really anything new here; these are the same properties that you have used for real numbers for years. As exemplified in Problem 6.4, we can prove these properties for complex numbers by using the corresponding properties for real numbers together with the definitions of addition and multiplication of complex numbers.

Problem 6.5:

(a) Find all values of k such that the product $(3 + 2i)(3 + ki)$ is a real number.

(b) Let $z = a + bi$, where a and b are real numbers, and let $\overline{z} = a - bi$. Prove that $z \cdot \overline{z}$ is real.

Solution for Problem 6.5:

(a) Expanding the given product, we have $(3 + 2i)(3 + ki) = 9 - 2k + (6 + 3k)i$. The result is only a real number when $6 + 3k = 0$, so $k = -2$.

(b) In part (a) we saw that the product of the nonreal numbers $3 + 2i$ and $3 - 2i$ is real. Notice that the only difference between these two numbers is the sign of the imaginary part. Let's see if we can generalize this observation by checking whether or not the product $(a + bi)(a - bi)$ is a real number. We find

$$(a + bi)(a - bi) = a(a - bi) + bi(a - bi) = a^2 - abi + abi - b^2i^2 = a^2 + b^2,$$

which is real because a and b are real.

□

Definition: If $z = a + bi$, then $a - bi$ is called the **conjugate** of z, denoted $\overline{z}$. So, we have $\overline{a + bi} = a - bi$. Together, a nonreal number and its conjugate are referred to as a **conjugate pair.**

Important: The product of a complex number and its conjugate is a real number.

Note that in order to prove this fact in Problem 6.5, we wrote a complex number in terms of its real and imaginary parts by writing $z = a + bi$. We also used this strategy in Problem 6.4.

Concept: If you are given a problem involving a general complex number z, it may help to write z in the form $a + bi$, and then express the problem in terms of a and b.

We've covered addition, subtraction, and multiplication of complex numbers. Let's put conjugates to work by turning to division.

Problem 6.6: Write $\dfrac{2-i}{3+2i}$ as a complex number. In other words, find real numbers a and b such that $\dfrac{2-i}{3+2i} = a + bi$.

Solution for Problem 6.6: Knowing that $(3 + 2i)(3 - 2i)$ is real, we multiply the given fraction by $\frac{3-2i}{3-2i}$, which equals 1, to produce an equivalent fraction with a real denominator:

$$\frac{2-i}{3+2i} = \frac{2-i}{3+2i} \cdot \frac{3-2i}{3-2i} = \frac{(2-i)(3-2i)}{(3+2i)(3-2i)} = \frac{6-4i-3i+2i^2}{3^2+2^2} = \frac{4-7i}{13}.$$

Therefore, we have $\dfrac{2-i}{3+2i} = \dfrac{4}{13} - \dfrac{7}{13}i$. □

The fact that $z\bar{z}$ is real for all complex numbers z gives us a way to divide by a complex number.

Important: If w and z are complex numbers and $z \neq 0$, then we express the quotient w/z as a complex number by multiplying both the numerator and denominator by the conjugate of the denominator, $\bar{z}$:

$$\frac{w}{z} = \frac{w}{z} \cdot \frac{\bar{z}}{\bar{z}} = \frac{w\bar{z}}{z\bar{z}}.$$

Since $z\bar{z}$ is always real, we can express $w\bar{z}/(z\bar{z})$ as a complex number, which means we can write it in the form $a + bi$ for some pair of real numbers a and b.

Problem 6.7: The **multiplicative inverse** of a number z is the number w such that $wz = 1$. For example, 0.25 is the multiplicative inverse of 4. Show that any nonzero complex number $a + bi$ has a multiplicative inverse, and express that inverse in terms of a and b.

Solution for Problem 6.7: We expect that the multiplicative inverse is simply $\frac{1}{z}$, expressed as a complex number using the process we learned in Problem 6.6. If we let

$$w = \frac{1}{a+bi} = \frac{1}{a+bi} \cdot \frac{a-bi}{a-bi} = \frac{a-bi}{a^2+b^2},$$

then

$$wz = \frac{a-bi}{a^2+b^2} \cdot (a+bi) = \frac{(a-bi)(a+bi)}{a^2+b^2} = \frac{a^2+b^2}{a^2+b^2} = 1.$$

This only fails if $a^2 + b^2 = 0$. Since a and b are real, their squares are nonnegative, and can only both be 0 if $a = b = 0$. Therefore, if $a + bi$ is a nonzero complex number, then its multiplicative inverse is $\frac{a-bi}{a^2+b^2}$. □

Problem 6.8: Find all complex numbers z such that $\dfrac{3z-5i}{2z+3} = 3 - 2i$.

Solution for Problem 6.8: We get rid of the fraction by multiplying both sides of the equation by the denominator of the fraction. This gives

$$3z - 5i = (2z + 3)(3 - 2i).$$

Expanding the right side gives

$$3z - 5i = 6z - 4iz + 9 - 6i.$$

Putting all the terms with z on the left and the terms without z on the right gives $-3z + 4iz = 9 - i$. Factoring z out of the left side gives $z(-3 + 4i) = 9 - i$. Dividing both sides by $-3 + 4i$ gives

$$z = \frac{9 - i}{-3 + 4i} = \frac{9 - i}{-3 + 4i} \cdot \frac{-3 - 4i}{-3 - 4i} = \frac{-31 - 33i}{25} = -\frac{31}{25} - \frac{33}{25}i.$$

□

Exercises

6.1.1 Find all z such that $4z^2 + 12 = 0$.

6.1.2 Let $a = 3 + 4i$ and $b = 12 - 5i$. Compute each of the following:

(a) $a - b$

(b) ab

(c) $a^2 + 3a + 2$

(d) $\frac{a}{\overline{b}} + \frac{\overline{a}}{b}$

6.1.3 Write each of the following in the form $a + bi$, where a and b are real:

(a) $\overline{2i} + \overline{7 - 2i}$

(b) $\dfrac{1}{\overline{2 + 3i} + \overline{1 + 2i}}$

6.1.4 Let $f(x) = \dfrac{x^6 + x^4}{x + 1}$. Find each of the following:

(a) $f(i)$

(b) $f(-i)$

(c) $f(i - 1)$

6.1.5 What is $(i - i^{-1})^{-1}$? *(Source: AHSME)*

6.1.6 Solve each of the following equations for z:

(a) $\dfrac{z + 3i}{z - 3} = 2$

(b) $\dfrac{1 + 2i}{3z} = 4 + 5i$

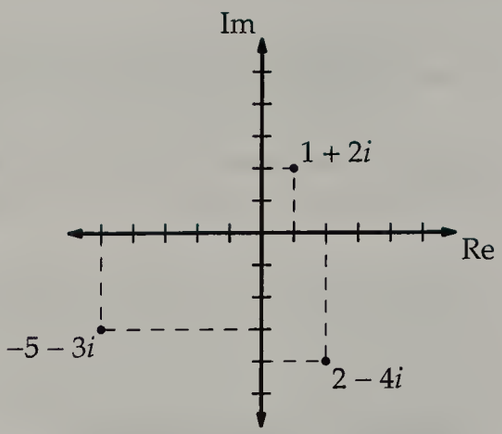
(c) $3z + \dfrac{z}{1 + i} = 10 - 4i$

6.1.7 Use the definitions of complex number addition and complex number multiplication to prove that complex numbers are distributive. In other words, prove that $w(z_1 + z_2) = wz_1 + wz_2$ for any complex numbers w, z_1, and z_2.

6.1.8 Let $S = i^n + i^{-n}$, where n is an integer. What are the possible values of S? *(Source: AHSME)*

6.2 The Complex Plane

A complex number can be represented as a point in the **complex plane**. Like the Cartesian plane, the complex plane has two axes: a horizontal **real axis** for the real part, and a vertical **imaginary axis** for the imaginary part. The real axis of the complex plane is labeled Re and the imaginary axis is labeled Im. Plotting the complex number $x + yi$ on the complex plane is the same as plotting the point (x, y) in the Cartesian plane. The figure at right shows the numbers $1 + 2i$, $-5 - 3i$, and $2 - 4i$ plotted in the complex plane.



Just as the point $(0, 0)$ is called the origin on the Cartesian plane, the point that represents 0 on the complex plane is also called the **origin** of the complex plane.

Problems

Problem 6.9:

(a) Let $w = 2 + 3i$. Plot w, $\overline{w}$, and $-w$ on the complex plane.

(b) For any complex number z, how is z related to $\overline{z}$ on the complex plane? How is z related to $-z$ on the complex plane?

Problem 6.10:

(a) Let $w = 2 + 4i$ and $z = 3 - 2i$. Plot w, z, and $w + z$ on the complex plane. Connect the origin to w, then connect w to $w + z$, then connect $w + z$ to z, then connect z to the origin. Notice anything interesting?

(b) Suppose w and z are nonzero complex numbers such that w/z is not real. Connect the origin to w, then connect w to $w + z$, then connect $w + z$ to z, then connect z to the origin. What type of quadrilateral must result? Why?

Problem 6.11:

(a) What is the distance between $3 + 4i$ and the origin in the complex plane?

(b) What is the distance between $a + bi$ and the origin in the complex plane?

Problem 6.12: Let $|a + bi| = \sqrt{a^2 + b^2}$.

(a) Evaluate $|5 - 12i|$ and $|3 - 3i|$.

(b) Suppose $w = 3 - 5i$ and $z = -2 + 7i$. Find $|w - z|$. Find the distance between w and z on the complex plane. Notice anything interesting?

(c) For any complex numbers w and z, how is $|w - z|$ related to the distance between w and z on the complex plane?

(d) Show that if $|z| = 0$, then $z = 0$.

Problem 6.13: Describe the graphs of each of the following:

(a) $|z| = 3$

(b) $|z - 4| = 3$

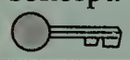
(c) $|z + 5 - 4i| = 2\sqrt{2}$

(d) $|z - w| = c$, where w is a complex number and c is real.

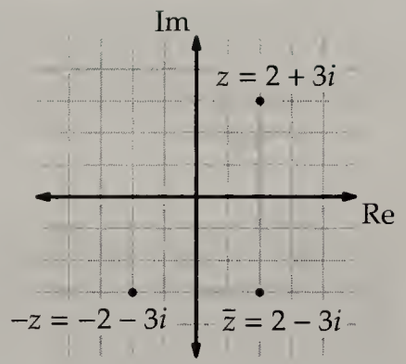
Problem 6.9: For any complex number z, how is z related to $\bar{z}$ on the complex plane? How is z related to $-z$ on the complex plane?

Solution for Problem 6.9: We start by experimenting with a specific value of z.

Concept: Trying specific examples can be a good way to discover general relationships.



Suppose $z = 2 + 3i$. Then, we have $\bar{z} = 2 - 3i$ and $-z = -2 - 3i$. We see that $\bar{z}$ is the reflection of z over the real axis and $-z$ is the reflection of $\bar{z}$ over the imaginary axis. Is this just a coincidence?



To see if this is true in general, we let $z = a + bi$. Then, we have $\bar{z} = a - bi$ and $-z = -a - bi$. So, z and $\bar{z}$ have the same real part and opposite imaginary parts. Therefore, z and $\bar{z}$ are reflections of each other over the real axis. Similarly, $\bar{z}$ and $-z$ have the same imaginary part, but opposite real parts, so $\bar{z}$ and $-z$ are reflections of each other over the imaginary axis.

But how are z and $-z$ related? Because $z = a + bi$ and $-z = -a - bi$, the points corresponding to z and $-z$ are in exactly opposite directions from the origin. We also see that z and $-z$ are the same distance from the origin. Therefore, $-z$ is the 180° rotation of z around the origin on the complex plane. (Note: We sometimes refer to the 180° rotation of a point A around a point P as the "reflection of point A through point P.") □

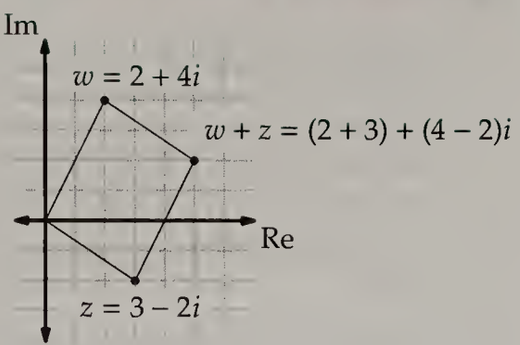
Important: In the complex plane, we have the following:

- The point $-z$ is a 180° rotation of point z about the origin.
- The points z and $\bar{z}$ are reflections of each other over the real axis.
- The points $\bar{z}$ and $-z$ are reflections of each other over the imaginary axis.

Don't memorize these relationships! If you understand why they are true, they should become obvious to you.

Problem 6.10: Suppose w and z are nonzero complex numbers such that w/z is not a real constant. (In other words, there is no real number a such that $w = az$.) Connect the origin to w, then connect w to $w + z$, then connect $w + z$ to z, then connect z to the origin. What type of quadrilateral must result? Why?

Solution for Problem 6.10: Again, we start with an example. Suppose $w = 2 + 4i$ and $z = 3 - 2i$, so $w + z = 5 + 2i$. In the diagram at right, we connect the points as suggested. We appear to have a parallelogram.



To see if this is always the case, we let $w = a + bi$ and $z = c + di$. So, we have $w + z = (a + c) + (b + d)i$. If $a \neq 0$, then the slope between w and the origin is b/a, as is the slope between $w + z$ and z. If $a = 0$, then both of these lines are vertical lines. In either case, these two sides of the quadrilateral are parallel. Similarly, the slopes of the other two sides are either both equal to d/c or they are both undefined (when both sides are vertical). So, these two sides are parallel as well. Therefore, both pairs of opposite sides of the quadrilateral are parallel, which means the quadrilateral is a parallelogram. □

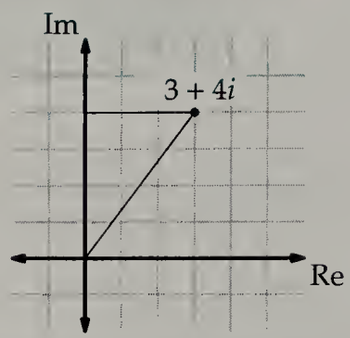
Problem 6.11:

(a) What is the distance between $3 + 4i$ and the origin in the complex plane?

(b) What is the distance between $a + bi$ and the origin in the complex plane?

Solution for Problem 6.11:

(a) Finding the distance between $3 + 4i$ and the origin on the complex plane is the same as finding the distance between $(3, 4)$ and $(0, 0)$ on the Cartesian plane. We can use the distance formula, or build a right triangle as shown at right, to find that the desired distance is $\sqrt{(3-0)^2 + (4-0)^2} = 5$.



(b) The distance between $a + bi$ and 0 on the complex plane is the same as the distance between (a, b) and $(0, 0)$ on the Cartesian plane. So, the desired distance is $\sqrt{a^2 + b^2}$.

□

Just as we use $|x|$ to denote the distance between a real number x and 0 on the number line, we use $|z|$ to represent the distance between a complex number z and the origin on the complex plane. We call $|z|$ the **magnitude** of z. The magnitude of a complex number is sometimes called the **norm** of the number.

Important: The **magnitude** of z, denoted by $|z|$, equals the distance from z to the origin on the complex plane. If $z = a + bi$, we have

$$|z| = \sqrt{a^2 + b^2}.$$

Problem 6.12:

(a) Evaluate $|5 - 12i|$ and $|3 - 3i|$.

(b) Suppose $w = 3 - 5i$ and $z = -2 + 7i$. Find $|w - z|$. Find the distance between w and z on the complex plane. Notice anything interesting?

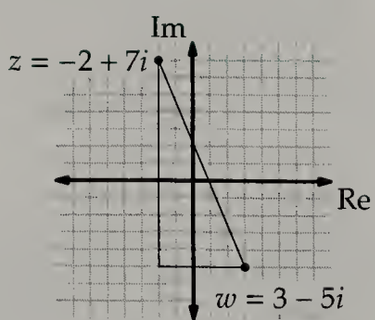
(c) For any complex numbers w and z, how is $|w - z|$ related to the distance between w and z on the complex plane?

(d) Show that if $|z| = 0$, then $z = 0$.

Solution for Problem 6.12:

(a) We have $|5 - 12i| = \sqrt{5^2 + (-12)^2} = 13$ and $|3 - 3i| = \sqrt{3^2 + (-3)^2} = 3\sqrt{2}$.

(b) We have $w - z = 5 - 12i$, and in the previous part we found that $|5 - 12i| = 13$. We plot w and z in the diagram at right. We see that the horizontal distance between them is 5 and the vertical distance between them is 12, so the segment connecting the two points is the hypotenuse of a right triangle with legs of lengths 5 and 12. Therefore, the Pythagorean Theorem tells us that the distance between w and z on the complex plane is $\sqrt{5^2 + 12^2} = 13$.



Notice that the distance between w and z equals $|w - z|$. Is this a coincidence?

(c) No, it's not a coincidence. Suppose $w = a + bi$ and $z = c + di$. Finding the distance between w and z in the complex plane is the same as finding the distance between (a, b) and (c, d) on the Cartesian plane. So, the distance between w and z is $\sqrt{(a-c)^2 + (b-d)^2}$.

Because $w - z = (a + bi) - (c + di) = (a - c) + (b - d)i$, we have

$$|w - z| = \sqrt{(a - c)^2 + (b - d)^2}.$$

This tells us that:

Important: The distance between complex numbers w and z on the complex plane is $|w - z|$.

(d) Geometrically, the equation $|z - 0| = 0$ tells us that the distance from z to 0 in the complex plane is 0. Therefore, we must have $z = 0$.

Algebraically, if we let $z = a + bi$, where a and b are real, then the equation $|z| = 0$ tells us that $\sqrt{a^2 + b^2} = 0$. Squaring both sides gives $a^2 + b^2 = 0$. Since a and b are real, their squares are nonnegative. So, we only have $a^2 + b^2 = 0$ if $a^2 = b^2 = 0$, which means $a = b = 0$. Therefore, we again have $z = 0$.

□

Problem 6.13: Describe the graphs of each of the following:

(a) $|z| = 3$

(b) $|z - 4| = 3$

(c) $|z + 5 - 4i| = 2\sqrt{2}$

(d) $|z - w| = c$, where w is a complex number and c is real.

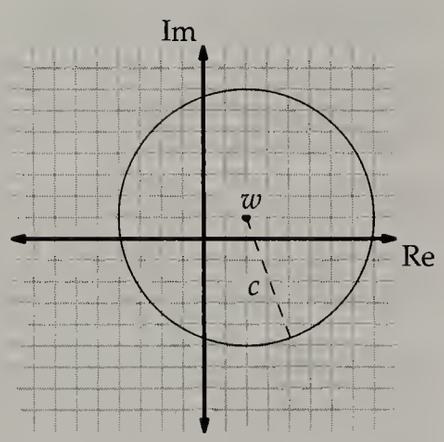
Solution for Problem 6.13:

(a) If $|z| = 3$, then the distance from z to the origin of the complex plane is 3. Therefore, z must be on the circle centered at the origin with radius 3. Moreover, every point z on this circle satisfies the equation $|z| = 3$, because every point on this circle is 3 units from the origin. Therefore, the graph $|z| = 3$ describes a circle of radius 3 centered at the origin.

(b) The equation $|z - 4| = 3$ tells us that the distance between z and 4 in the complex plane is 3, so the graph is a circle of radius 3 centered at the point that represents 4. Since $4 = 4 + 0i$, the center of the circle is on the real axis, 4 units to the right of the origin.

(c) Writing $z + 5 - 4i$ as $z - (-5 + 4i)$, we see that the graph of the equation $|z - (-5 + 4i)| = 2\sqrt{2}$ consists of all points that are a distance of $2\sqrt{2}$ from the point $-5 + 4i$. Therefore, the graph is a circle of radius $2\sqrt{2}$ centered at $-5 + 4i$.

(d) Generalizing our work from parts (a)-(c), we see that the graph of the equation $|z - w| = c$ consists of all points that are a distance c from the complex number w. In other words, the graph is a circle of radius c centered at w (shown at right for $w = 2 + i$ and $c = 6$).



□

Exercises

6.2.1 Plot each of the following in the complex plane:

(a) $4 + 7i$ (b) $-6 - 2i$ (c) $(3 + i)(-2 + 5i)$

6.2.2 Find the magnitude of each of the following complex numbers:

(a) $24 - 7i$

(b) $2 + 2i\sqrt{3}$

(c) $(1 + 2i)(2 + i)$

6.2.3 Let $w = 3 + 5i$ and $z = 12 + 2i$. Find the area of the convex quadrilateral in the complex plane that has vertices w, z, $\overline{w}$, and $\overline{z}$.

6.2.4 Find the distance between the points $4 + 7i$ and $-3 - 17i$ in the complex plane.

6.2.5 Show that the midpoint of the segment connecting z_1 and z_2 on the complex plane is $(z_1 + z_2)/2$.

6.2.6

(a) Find the magnitude of $\dfrac{1+2i}{2+i}$.

(b) Find the magnitude of $\dfrac{6+11i}{11+6i}$. (You can use a calculator for this part.)

(c) Notice anything interesting? Can you generalize your observations from the first two parts?

6.2.7 Four complex numbers lie at the vertices of a square in the complex plane. Three of the numbers are $1 + 2i$, $-2 + i$ and $-1 - 2i$. What is the fourth number? *(Source: AMC 12)*

6.3 Real and Imaginary Parts

Not only can we produce a geometric representation of complex numbers by considering the real and imaginary parts of the numbers, but we can also use the real and imaginary parts of a complex numbers to prove important relationships and solve a variety of problems.

We often use Re(z) to refer to the real part of the complex number z, and we use Im(z) to refer to the imaginary part of z. So, for example $\text{Re}(2 - 4i) = 2$ and $\text{Im}(2 - 4i) = -4$.

Problems

Problem 6.14: Show that $\text{Re}(z) = \dfrac{z+\overline{z}}{2}$ and $\text{Im}(z) = \dfrac{z-\overline{z}}{2i}$.

Problem 6.15: Let z and w be complex numbers.

(a) Show that $\overline{z+w} = \overline{z} + \overline{w}$.

(b) Show that $\overline{zw} = \overline{z} \cdot \overline{w}$.

Problem 6.16:

(a) Show that $\overline{\overline{z}} = z$ for all complex numbers z.

(b) Show that $\overline{z} = z$ if and only if z is real.

(c) Show that $\overline{z} = -z$ if and only if z is imaginary.

Problem 6.17: Solve the equation $z + 2\overline{z} = 6 - 4i$ for z.

Problem 6.18:

(a) Prove that $z\bar{z} = |z|^2$ for all complex numbers z.

(b) Prove that $|zw| = |z||w|$ for all complex numbers z and w.

Problem 6.19: Suppose $|z + w| = 2$ and $|z| = |w| = 3$. Find $|z - w|$.

Problem 6.20: In this problem, we find all complex numbers z such that $z^2 = 21 - 20i$.

(a) Let $z = a + bi$ in the given equation. Find a system of equations involving a and b.

(b) Solve for b in terms of a in one of the equations, and substitute the expression you found for b into the other equation.

(c) Solve the equation you formed in part (b) for all possible values of a.

(d) Find all complex numbers z such that $z^2 = 21 - 20i$.

Problem 6.21: Find and graph all z such that $|z - 3| = |z + 2i|$.

Problem 6.22: In this problem, we find $|z|$ if the complex number z satisfies $z + |z| = 2 + 8i$.

(a) Let $r = |z|$. Express z in terms of r.

(b) What is the real part of z? What is the imaginary part of z? Find $|z|$.

(c) Solve the problem again by letting $z = a + bi$.

(Source: AHSME)

Problem 6.14: Show that $\text{Re}(z) = \dfrac{z + \bar{z}}{2}$ and $\text{Im}(z) = \dfrac{z - \bar{z}}{2i}$.

Solution for Problem 6.14: Let $z = a + bi$ so that $\bar{z} = a - bi$. Then

$$\frac{z + \bar{z}}{2} = \frac{(a + bi) + (a - bi)}{2} = \frac{2a}{2} = a,$$

$$\frac{z - \bar{z}}{2i} = \frac{(a + bi) - (a - bi)}{2i} = \frac{2bi}{2i} = b.$$

So, $\text{Re}(z) = a = \dfrac{z + \bar{z}}{2}$ and $\text{Im}(z) = b = \dfrac{z - \bar{z}}{2i}$. □

Important: For any complex number z, we have

$$\text{Re}(z) = \frac{z + \bar{z}}{2} \quad \text{and} \quad \text{Im}(z) = \frac{z - \bar{z}}{2i}.$$

Problem 6.15: Let z and w be complex numbers.

(a) Show that $\overline{z + w} = \bar{z} + \bar{w}$.

(b) Show that $\overline{zw} = \bar{z} \cdot \bar{w}$.

Solution for Problem 6.15: Let $z = a + bi$ and $w = c + di$, where a, b, c, and d are real numbers.

(a) Since $z + w = a + bi + c + di = (a + c) + (b + d)i$, the conjugate of $z + w$ is

$$\overline{z + w} = \overline{(a + c) + (b + d)i} = (a + c) - (b + d)i.$$

Similarly, the conjugate of z is $\overline{z} = a - bi$, and the conjugate of w is $\overline{w} = c - di$, and their sum is

$$a - bi + c - di = (a + c) - (b + d)i.$$

So, we have $\overline{z + w} = \overline{z} + \overline{w}$.

(b) We have

$$zw = (a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i,$$

so the conjugate of zw is $\overline{zw} = (ac - bd) - (ad + bc)i$. Also, we have

$$\overline{z} \cdot \overline{w} = (a - bi)(c - di) = ac - adi - bci + bdi^2 = (ac - bd) - (ad + bc)i,$$

so $\overline{zw} = \overline{z} \cdot \overline{w}$.

□

Important: For any complex numbers z and w, we have $\overline{z + w} = \overline{z} + \overline{w}$ and $\overline{zw} = \overline{z} \cdot \overline{w}$.

We'll be seeing these two very important relationships again when we study nonreal roots of polynomials in Section 6.4.

Problem 6.16:

(a) Show that $\overline{\overline{z}} = z$ for all complex numbers z.

(b) Show that $\overline{z} = z$ if and only if z is real.

(c) Show that $\overline{z} = -z$ if and only if z is imaginary.

Solution for Problem 6.16: Let $z = a + bi$, where a and b are real numbers.

(a) Since $\overline{z} = a - bi$, we have $\overline{\overline{z}} = \overline{a - bi} = a + bi = z$.

(b) Since $\overline{z} = a - bi$, the equation $\overline{z} = z$ becomes $a - bi = a + bi$. Subtracting a from both sides gives $-bi = bi$. Therefore, we have $2bi = 0$, so $b = 0$. This gives us $z = a + bi = a + 0i = a$, so z is real. This shows that $\overline{z} = z$ only if z is real.

We're not finished with our proof! We have only shown that when $\overline{z} = z$, then z is real. (In other words, we have $\overline{z} = z$ *only if* z is real.) We have not shown that the equation $\overline{z} = z$ is true for all real numbers. In other words, we haven't shown that $\overline{z} = z$ *if* z is real.

To show that $\overline{z} = z$ whenever z is real, we let $z = a$, where a is a real number. We then have $\overline{z} = \overline{a} = a$, so $\overline{z} = z$ if z is real.

Important: Suppose A and B are two mathematical statements. The statement "A if and only if B" means both of the following:

- If A is true, then B is also true.
- If B is true, then A is also true.

Therefore, as we just saw, proving an "if and only if" statement requires proving two statements, not just one.

(c) Since $\overline{z} = a - bi$, the equation $\overline{z} = -z$ becomes $a - bi = -a - bi$. Adding bi to both sides gives $a = -a$, so $a = 0$. Therefore, we have $z = a + bi = bi$, so z is imaginary. This shows that $\overline{z} = -z$ only if z is imaginary.

As with part (b), this is an "if and only if" statement. We must also show that $\overline{z} = -z$ if z is imaginary. Letting $z = bi$, we have $\overline{z} = \overline{bi} = -bi = -z$, as desired.

□

Problem 6.17: Solve $z + 2\overline{z} = 6 - 4i$ for z.

Solution for Problem 6.17: Letting $z = a + bi$, for some real numbers a and b, the given equation becomes $(a + bi) + 2(a - bi) = 6 - 4i$. Simplifying the left side gives $3a - bi = 6 - 4i$. Since two complex numbers are equal if and only if their real parts are equal and their imaginary parts are equal, we have $3a = 6$ and $-b = -4$. Solving these two equations gives $a = 2$ and $b = 4$, so $z = a + bi = 2 + 4i$. □

Problem 6.18:

(a) Prove that $z\overline{z} = |z|^2$ for all complex numbers z.

(b) Prove that $|zw| = |z||w|$ for all complex numbers z and w.

Solution for Problem 6.18: Let $z = a + bi$ and $w = c + di$ for real numbers a, b, c, and d.

(a) Since $\overline{z} = a - bi$, we have $z\overline{z} = (a + bi)(a - bi) = a^2 + b^2 = \left(\sqrt{a^2 + b^2}\right)^2 = |a + bi|^2 = |z|^2$.

(b) First, we note that $zw = (a + bi)(c + di) = ac - bd + (ad + bc)i$. So, we have

$$\begin{aligned}
|zw| = |ac - bd + (ad + bc)i| &= \sqrt{(ac - bd)^2 + (ad + bc)^2} \\
&= \sqrt{a^2c^2 - 2abcd + b^2d^2 + a^2d^2 + 2abcd + b^2c^2} \\
&= \sqrt{a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2} \\
&= \sqrt{a^2(c^2 + d^2) + b^2(c^2 + d^2)} \\
&= \sqrt{(a^2 + b^2)(c^2 + d^2)} \\
&= \sqrt{a^2 + b^2}\sqrt{c^2 + d^2} \\
&= |a + bi||c + di| = |z||w|.
\end{aligned}$$

We could also have used part (a) together with a little cleverness to solve part (b). From part (a), we have

$$|zw|^2 = zw(\overline{zw}) = zw\overline{z} \cdot \overline{w} = (z\overline{z})(w\overline{w}) = |z|^2|w|^2.$$

Since all the magnitudes are nonnegative numbers, we can take the positive square root of both sides to get $|zw| = |z||w|$.

□

Important: The relationship $|z|^2 = z\overline{z}$ can be very helpful in problems involving the magnitudes of complex numbers.

Problem 6.19: Suppose $|z + w| = 2$ and $|z| = |w| = 3$. Find $|z - w|$.

Solution for Problem 6.19: There's no clear way to express $|z-w|$ in terms of $|z+w|$, $|z|$, and $|w|$, so we write it in terms of z, w, $\overline{z}$, and $\overline{w}$:

$$|z-w|^2 = (z-w)\overline{(z-w)} = (z-w)(\overline{z}-\overline{w}) = z\overline{z} - z\overline{w} - w\overline{z} + w\overline{w}.$$

Notice that we work with $|z-w|^2$ rather than $|z-w|$ to avoid square roots.

Concept: It's sometimes easier to work with the square of the magnitude of a complex number than with the magnitude itself.

From $|z| = |w| = 3$, we find that $z\overline{z} = |z|^2 = 9$ and $w\overline{w} = |w|^2 = 9$, so we have

$$|z-w|^2 = z\overline{z} - z\overline{w} - w\overline{z} + w\overline{w} = 18 - z\overline{w} - w\overline{z}.$$

We have one more magnitude to explore. As with the others, we start by looking at its square. We have $|z+w|^2 = 4$ and

$$|z+w|^2 = (z+w)\overline{(z+w)} = z\overline{z} + z\overline{w} + w\overline{z} + w\overline{w} = 9 + z\overline{w} + w\overline{z} + 9,$$

so $18 + z\overline{w} + w\overline{z} = 4$, which means $z\overline{w} + w\overline{z} = -14$. Therefore, we have $|z-w|^2 = 18 - (z\overline{w} + w\overline{z}) = 18 - (-14) = 32$, from which we find $|z-w| = 4\sqrt{2}$. □

Problem 6.20: Find all complex numbers z such that $z^2 = 21 - 20i$.

Solution for Problem 6.20: We might start by taking the square root of both sides, so that $z = \pm\sqrt{21-20i}$. But what does $\sqrt{21-20i}$ really mean? The real-valued function $f(x) = \sqrt{x}$ is only defined for $x \geq 0$. Unfortunately, this doesn't help us define the square root of a nonreal number. Instead of trying to define $\sqrt{21-20i}$, let's look for complex numbers whose squares equal $21 - 20i$.

Let $z = a + bi$, so we have $(a+bi)^2 = 21 - 20i$. Expanding the left side gives

$$a^2 - b^2 + 2abi = 21 - 20i.$$

Since a and b are real, we can equate the real parts of both sides to get $a^2 - b^2 = 21$, and equate the imaginary parts to find $2ab = -20$. Solving $2ab = -20$ for b gives $b = -10/a$. Substituting this into $a^2 - b^2 = 21$ gives

$$a^2 - \frac{100}{a^2} = 21.$$

Multiplying this by a^2 gives us $a^4 - 21a^2 - 100 = 0$. Factoring gives $(a^2-25)(a^2+4) = 0$. However, a is real, so a^2 must be nonnegative. Therefore, we have $a^2 = 25$, so $a = \pm 5$. Since $b = -10/a$ we have $b = -2$ when $a = 5$, and $b = 2$ when $a = -5$. So, the two possible values of z are $\pm(5-2i)$. □

Having found the two complex numbers whose squares equal $21 - 20i$, we can see some of the difficulty that arises from defining $\sqrt{21-20i}$. Which of the values $5-2i$ or $-5+2i$ should we take to be the square root? We can't simply call one "positive" and one "negative."

Due to these difficulties, we try to avoid using expressions like $\sqrt{21-20i}$, and instead think of the "numbers whose squares equal $21 - 20i$." We may sometimes refer to these two numbers as "the square roots" of $21 - 20i$, and may even refer to one or the other as "a square root" of $21 - 20i$. But we generally won't choose one or the other to equal $\sqrt{21-20i}$ the way we choose 5 to equal $\sqrt{25}$.

We have these same difficulties when trying to define $\sqrt{-81}$. For example, which of $9i$ or $-9i$ is "the" square root of -81? However, as we saw in Problem 6.1, we may occasionally stumble on expressions like $\sqrt{-81}$ when

solving equations like $z^2 = -81$. When this happens, the square root should always be preceded by $\pm$, as in $z = \pm\sqrt{-81}$. In other words, we are referring to both "square roots" of -81; that is, we are considering both numbers whose squares are -81.

Problem 6.21: Find and graph all z such that $|z - 3| = |z + 2i|$.

Solution for Problem 6.21: Solution 1: Algebra. Letting $z = a + bi$, for real numbers a and b, the equation becomes $|a + bi - 3| = |a + bi + 2i|$, or

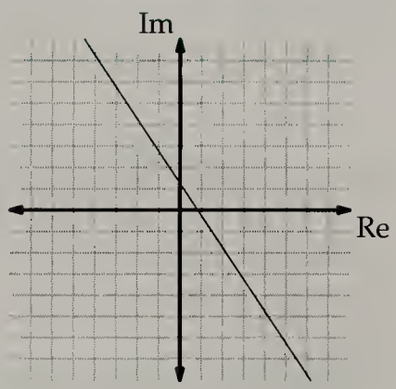
$$|(a - 3) + bi| = |a + (b + 2)i|.$$

Applying the definition of the magnitude of a complex number to both sides of this equation, we have

$$\sqrt{(a - 3)^2 + b^2} = \sqrt{a^2 + (b + 2)^2}.$$

Squaring both sides of this equation and expanding the squares of binomials gives

$$a^2 - 6a + 9 + b^2 = a^2 + b^2 + 4b + 4.$$



Simplifying this equation gives $6a + 4b = 5$, so our equation is $6 \cdot \text{Re}(z) + 4 \cdot \text{Im}(z) = 5$. The graph of this equation is a line, shown at right above. We can write the equation $6 \cdot \text{Re}(z) + 4 \cdot \text{Im}(z) = 5$ in terms of z and $\bar{z}$ by noting that $\text{Re}(z) = (z+\bar{z})/2$ and $\text{Im}(z) = (z-\bar{z})/(2i)$. Substituting these into $6 \cdot \text{Re}(z) + 4 \cdot \text{Im}(z) = 5$ gives us $(3-2i)z+(3+2i)\bar{z} = 5$.

We also could have tackled this problem algebraically by squaring $|z - 3| = |z + 2i|$ to get $|z - 3|^2 = |z + 2i|^2$. We then use the fact that $|w|^2 = w\bar{w}$ for all complex numbers w, which gives us $(z - 3)\overline{(z - 3)} = (z + 2i)\overline{(z + 2i)}$. Since $\overline{z - 3} = \bar{z} - \overline{3} = \bar{z} - 3$ and $\overline{(z + 2i)} = \bar{z} + \overline{2i} = \bar{z} - 2i$, we have $(z - 3)(\bar{z} - 3) = (z + 2i)(\bar{z} - 2i)$. Expanding both sides and rearranging gives $(3 - 2i)z + (3 + 2i)\bar{z} = 5$, as before.

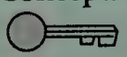
Solution 2: Geometric intuition. We rewrite the equation so that both magnitudes are of differences between complex numbers, $|z - 3| = |z - (-2i)|$. The quantity on the left side of the equation is the distance from z to 3 in the complex plane. Similarly, the right side is the distance from z to $-2i$ in the complex plane. Therefore, our equation tells us that z is equidistant from 3 and $-2i$ in the complex plane. So, z must be on the perpendicular bisector of the segment that connects 3 and $-2i$ in the complex plane.

The slope of the line connecting 3 and $-2i$ is $\frac{2}{3}$, so the slope of its perpendicular bisector is $-\frac{3}{2}$. The perpendicular bisector passes through the midpoint of 3 and $-2i$, which is $\frac{3-2i}{2} = \frac{3}{2} - i$. Using this point and the slope of the line, we can write a point-slope form of the line:

$$\text{Im}(z) - (-1) = -\frac{3}{2}\left(\text{Re}(z) - \frac{3}{2}\right).$$

Rearranging this equation gives us $6 \cdot \text{Re}(z) + 4 \cdot \text{Im}(z) = 5$, as before. □

The key to our second solution is using the geometric interpretations of $|z - 3|$ and $|z + 2i|$.

Concept:  Many complex number concepts have a strong geometric significance as well as an algebraic one, so if you have trouble understanding an algebraic relationship regarding complex numbers, try considering it in the complex plane.

Problem 6.22: The complex number z satisfies $z + |z| = 2 + 8i$. What is $|z|$? *(Source: AHSME)*

Solution for Problem 6.22: Solution 1: Keep your eye on the ball. We want $|z|$, and our equation involves $|z|$, so we let $|z| = r$. Then, we have $z = 2 + 8i - |z| = (2 - r) + 8i$. Written this way, we can see that the complex number z has real part $2 - r$ and imaginary part 8. Therefore, we have

$$|z|^2 = (2 - r)^2 + 8^2 = r^2 - 4r + 68.$$

We also have $|z|^2 = r^2$, and combining this with our equation above gives $r^2 = r^2 - 4r + 68$. Solving this equation gives us $r = 17$, so $|z| = 17$.

Solution 2: Let $z = a + bi$. Then, the given equation is

$$a + bi + \sqrt{a^2 + b^2} = 2 + 8i.$$

The only imaginary term on the left is bi, so we have $b = 8$. Equating the real parts of both sides then gives $a + \sqrt{a^2 + 64} = 2$. Isolating the radical gives $\sqrt{a^2 + 64} = 2 - a$ and squaring both sides gives $a^2 + 64 = 4 - 4a + a^2$. Solving this equation gives us $a = -15$. Because we squared the equation $\sqrt{a^2 + 64} = 2 - a$, we must check if the solution is extraneous. It isn't, so $|z| = |-15 + 8i| = 17$. □

Exercises

6.3.1 Find $|(10 + 24i)(8 - 6i)|$ without finding the product $(10 + 24i)(8 - 6i)$.

6.3.2 Show that $|z| = |\overline{z}|$.

6.3.3 Solve $3z + 4\overline{z} = 12 - 5i$ for z.

6.3.4 Solve each of the following equations:

(a) $z^2 = 2i$ (b) $z^2 = -5 + 12i$ (c) $z^2 = 24 - 10i$

6.3.5 Let $z = 3 + 4i$ and $w = 5 - 12i$. Evaluate the following expressions:

(a) $\left|\frac{1}{z}\right|$ (b) $\frac{1}{|z|}$ (c) $\left|\frac{z}{w}\right|$ (d) $\frac{|z|}{|w|}$

Compare your answers in parts (a) and (b), and compare your answers in parts (c) and (d). Notice anything interesting? (You should!) Prove that your observations will hold for any nonzero complex numbers.

6.3.6 Find all complex numbers z such that $z/\overline{z}$ is

(a) a real number (b) an imaginary number

6.3.7 Two solutions of $x^4 - 3x^3 + 5x^2 - 27x - 36 = 0$ are pure imaginary numbers. Find these two solutions. *(Source: ARML)* **Hints:** 174

6.4 Nonreal Roots of Polynomials

We can sometimes find roots of a polynomial by factoring the polynomial. For example, to find the roots of the quadratic $x^2 - 5x + 6$, we factor the quadratic as $(x - 2)(x - 3)$. Then, we have $(x - 2)(x - 3) = 0$ if and only if $x - 2 = 0$ or $x - 3 = 0$, which gives us the roots $x = 2$ and $x = 3$. So, we see that if we can factor a polynomial as the product of linear factors, then we can quickly find the roots of the polynomial, since each linear factor clearly has exactly one root.

We run into a complication with the polynomial $x^2 + 1$. We can't factor this as the product of two linear terms with real coefficients, and the polynomial doesn't have any real roots. So, we introduced the idea of complex

numbers to describe the roots of $x^2 + 1$, writing them as $x = i$ and $x = -i$, where i is the number whose square is -1. Complex numbers therefore also allow us to factor $x^2 + 1$ as a product of linear factors:

$$x^2 + 1 = (x - i)(x + i).$$

This might make you wonder if we need any more special types of numbers beyond complex numbers to factor more complicated polynomials as a product of linear factors. The answer turns out to be "no"; complex numbers are sufficient to factor any polynomial with complex coefficients. In order to see why, we'll need the **Fundamental Theorem of Algebra**.

Important: The **Fundamental Theorem of Algebra** states that every nonconstant one-variable polynomial with complex coefficients has at least one complex root.

Unfortunately, to prove the Fundamental Theorem of Algebra, we need tools that are much more advanced than those we will study in this text.

We'll also use the Factor Theorem and the Rational Root Theorem to study polynomials.

Important: The **Factor Theorem** states that for any complex number a, the binomial $x - a$ is a factor of the polynomial $f(x)$ if and only if $f(a) = 0$.

Important: The **Rational Root Theorem** states that if $\frac{p}{q}$ is a rational root of

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

where p, q, and all of the a_i are integers and $\gcd(p, q) = 1$, then a_0 is a multiple of p and a_n is a multiple of q.

The Factor Theorem and the Rational Root Theorem are covered in detail in Art of Problem Solving's *Intermediate Algebra*.

Problems

Problem 6.23: Show that if complex numbers w and z satisfy $wz = 0$, then either $w = 0$ or $z = 0$.

Problem 6.24: Consider the polynomial $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$, where $a_n, a_{n-1}, \ldots, a_0$ are complex numbers and $a_n \neq 0$.

(a) Show that there exist complex numbers $r_1, r_2, \ldots, r_n$ such that $f(x) = a_n(x - r_1)(x - r_2)\cdots(x - r_n)$.

(b) Use part (a) to explain why f has exactly n complex roots, where the same number may appear multiple times among the roots. (For example, the polynomial x^3 has the three roots 0, 0, and 0.)

Problem 6.25: Find the roots of the following polynomials:

(a) $f(x) = x^3 - 2x^2 + 3x - 18$

(b) $g(x) = 2x^4 - 8x^3 - 3x^2 + 22x - 40$

Problem 6.26: Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ and $q(x) = a_m x^m$ be polynomials with real coefficients.

(a) Show that $\overline{q(z)} = q(\overline{z})$ for all complex numbers z.

(b) Show that $\overline{p(z)} = p(\overline{z})$ for all complex numbers z.

(c) Show that if $p(z) = 0$, then $p(\overline{z}) = 0$.

Problem 6.27: Find the monic polynomial $f(x)$ of minimal degree such that $f(2-i) = f(1+4i) = 0$ and all the coefficients of $f(x)$ are real. (The **degree** of a polynomial $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$, with $a_n \neq 0$, is n, and the polynomial is called **monic** if $a_n = 1$. For example, $x^6 + 3x^2 + 7$ is a monic polynomial with degree 6, and $3x^5 - x + 1$ is a degree 5 polynomial that is not monic.)

Problem 6.28: Consider the equation $c_4 z^4 + ic_3 z^3 + c_2 z^2 + ic_1 z + c_0 = 0$, where c_0, c_1, c_2, c_3, and c_4 are real constants.

(a) Find a function h such that substituting $z = h(x)$ into the polynomial above produces a fourth degree polynomial in x with real coefficients.

(b) If one solution for z to the original equation is $a + bi$, where a and b are real, then what value of x solves your equation from part (a)?

(c) What other value of x must be a solution to your equation from (a)?

(d) Prove that $z = -a + bi$ is a solution to the original equation.

(Source: AHSME)

Problem 6.23: Show that if complex numbers w and z satisfy $wz = 0$, then either $w = 0$ or $z = 0$.

Solution for Problem 6.23: If w is 0, then the "$w = 0$ or $z = 0$" condition is satisfied. If w is not zero, then we can divide both sides by w to get $z = \frac{0}{w} = 0$, and again the "$w = 0$ or $z = 0$" condition is satisfied. □

We can extend the result of Problem 6.23 to any number of complex numbers: if $z_1 z_2 z_3 \cdots z_n = 0$, then at least one of the numbers $z_1, z_2, \ldots, z_n$ must be 0. This property is called the **zero-product property**. It's the same property that we use for real numbers to solve a quadratic equation like $x^2 - 5x + 6 = 0$. We factor the equation as $(x-2)(x-3) = 0$, and then note that if the product $(x-2)(x-3)$ is 0, then either $x - 2 = 0$ or $x - 3 = 0$. So, the solutions are $x = 2$ and $x = 3$. Of course, as we have seen, there are polynomials that cannot be factored as a product of linear terms with real coefficients. Let's see what happens if we allow complex coefficients in the factors.

Problem 6.24: Consider the polynomial $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$, where $a_n, a_{n-1}, \ldots, a_0$ are complex numbers and $a_n \neq 0$.

(a) Show that there exist complex numbers $r_1, r_2, \ldots, r_n$ such that $f(x) = a_n(x - r_1)(x - r_2)\cdots(x - r_n)$.

(b) Use part (a) to explain why f has exactly n complex roots, where the same number may appear multiple times among the roots. (For example, the polynomial x^3 has the three roots 0, 0, and 0.)

Solution for Problem 6.24:

(a) The Fundamental Theorem of Algebra tells us that this polynomial has a complex root, which we'll call r_1. So, if we divide $x - r_1$ into $f(x)$, the quotient is a polynomial $g(x)$ with degree $n - 1$. By the Factor Theorem, we therefore have

$$f(x) = (x - r_1)g(x) = (x - r_1)(b_{n-1}x^{n-1} + b_{n-2}x^{n-2} + \cdots + b_0)$$

for some constants $b_{n-1}, b_{n-2}, \ldots, b_0$. Applying the Fundamental Theorem of Algebra to $g(x)$, we know that it must have a complex root that we'll call r_2. Dividing $g(x)$ by $x - r_2$ gives a degree $n - 2$ quotient. Again

by the Factor Theorem, we have

$$f(x) = (x - r_1)(x - r_2)h(x)$$

for some degree $n - 2$ polynomial $h(x)$. Continuing in this manner, the quotient at each step has degree 1 less than the previous quotient. Therefore, we can express $f(x)$ as the product of a constant and n linear factors:

$$f(x) = a_n(x - r_1)(x - r_2)\cdots(x - r_n).$$

(b) If $f(x) = 0$, then product $(x - r_1)(x - r_2)\cdots(x - r_n)$ must equal 0. By the zero-product property, the product of several complex numbers equals 0 if and only if at least one of the numbers is 0. Therefore, we can only have $f(x) = 0$ if $x - r_i = 0$ for some i. The only solution to $x - r_i = 0$ is $x = r_i$, which is a complex number. This tells us that f has exactly n complex numbers as its roots.

□

Important: Every polynomial of degree n has exactly n complex roots. (A number may appear multiple times among the roots of the polynomial. For example, the polynomial $(x - 1)(x - 2)^2$ has the three complex roots 1, 2, and 2.)

Problem 6.25: Find the roots of the following polynomials:

(a) $f(x) = x^3 - 2x^2 + 3x - 18$

(b) $g(x) = 2x^4 - 8x^3 - 3x^2 + 22x - 40$

Solution for Problem 6.25:

(a) We use the Rational Root Theorem to hunt for integer roots by testing divisors of 18 to see if they are roots of $f(x)$. We find that $f(3) = 0$ and

$$f(x) = (x - 3)(x^2 + x + 6).$$

We can't factor $x^2 + x + 6$ easily, so we turn to the quadratic formula to find that the roots of $x^2 + x + 6$ are

$$x = \frac{-1 \pm i\sqrt{23}}{2}.$$

So, the three roots of $f(x)$ are 3 and $\dfrac{-1 \pm i\sqrt{23}}{2}$.

(b) We start by searching for rational roots. We find that $x + 2$ divides $g(x)$ evenly and gives us

$$g(x) = (x + 2)(2x^3 - 12x^2 + 21x - 20).$$

We continue by searching for roots of $2x^3 - 12x^2 + 21x - 20$. Again using the Rational Root Theorem to restrict our search, we find that $x - 4$ is a factor. Dividing $2x^3 - 12x^2 + 21x - 20$ by $x - 4$ gives us

$$g(x) = (x + 2)(x - 4)(2x^2 - 4x + 5).$$

As in part (a), we tackle the quadratic with the quadratic formula and find that its roots are

$$x = \frac{4 \pm 2i\sqrt{6}}{4} = 1 \pm \frac{\sqrt{6}}{2}i.$$

So, the four roots of $g(x)$ are -2, 4, $1 - \frac{\sqrt{6}}{2}i$, and $1 + \frac{\sqrt{6}}{2}i$.

□

For both polynomials in Problem 6.25, the nonreal roots come in conjugate pairs. In other words, if $a + bi$ is a root, then so is $a - bi$. From the quadratic formula, we can see that nonreal roots of a quadratic with real coefficients must form a conjugate pair. Let's see if we can prove a similar statement for higher degree polynomials.

Problem 6.26: Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ and $q(x) = a_m x^m$ be polynomials with real coefficients.

(a) Show that $\overline{q(z)} = q(\overline{z})$ for all complex numbers z.

(b) Show that $\overline{p(z)} = p(\overline{z})$ for all complex numbers z.

(c) Show that if $p(z) = 0$, then $p(\overline{z}) = 0$.

Solution for Problem 6.26:

(a) In Problem 6.15, we showed that for all complex numbers w and z, we have $\overline{wz} = \overline{w} \cdot \overline{z}$. Using this relationship, along with the fact that a_m is real, we see that

$$\overline{a_m z^m} = \overline{a_m} \cdot \overline{z^m} = a_m \overline{z^m}.$$

Now, we note that

$$\overline{z^m} = \overline{z} \cdot \overline{z^{m-1}} = \overline{z} \cdot \overline{z} \cdot \overline{z^{m-2}} = \cdots = \underbrace{\overline{z} \cdot \overline{z} \cdots \overline{z}}_{m\ \overline{z}\text{'s}} = \overline{z}^m,$$

which gives us $\overline{q(z)} = \overline{a_m z^m} = a_m \overline{z^m} = a_m \overline{z}^m = q(\overline{z})$.

(b) In Problem 6.15, we also showed that for all complex numbers w and z, we have $\overline{w+z} = \overline{w} + \overline{z}$. Combining this with part (a), we have

$$\begin{aligned}\overline{p(z)} &= \overline{a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0} \\ &= \overline{a_n z^n} + \overline{a_{n-1}z^{n-1}} + \cdots + \overline{a_1 z} + \overline{a_0} \\ &= a_n \overline{z}^n + a_{n-1}\overline{z}^{n-1} + \cdots + a_1 \overline{z} + a_0 \\ &= p(\overline{z}),\end{aligned}$$

as desired.

(c) From part (b), we have $p(\overline{z}) = \overline{p(z)}$. So, if $p(z) = 0$, then we have $p(\overline{z}) = \overline{p(z)} = \overline{0} = 0$.

□

Important: Let $p(x)$ be a polynomial with real coefficients. If z is a root of $p(x)$, then $\overline{z}$ is also a root of $p(x)$. In other words, the nonreal roots of $p(x)$ come in conjugate pairs.

We can sometimes use this fact to build a polynomial given some, but not all, of its roots.

Problem 6.27: Find the monic polynomial $f(x)$ of minimal degree such that $f(2 - i) = f(1 + 4i) = 0$ and all the coefficients of $f(x)$ are real. (The **degree** of a polynomial $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$, with $a_n \neq 0$, is n, and the polynomial is called **monic** if $a_n = 1$. For example, $x^6 + 3x^2 + 7$ is a monic polynomial with degree 6, and $3x^5 - x + 1$ is a degree 5 polynomial that is not monic.)

Solution for Problem 6.27: Since $2 - i$ is a root of f and the coefficients of f are real, $2 + i$ must also be a root of f. Therefore, the quadratic with roots $2 - i$ and $2 + i$ must be a factor of f. Here are two ways to find this quadratic:

Method 1: Multiply the linear factors. Since $2 - i$ and $2 + i$ are roots, the desired quadratic is

$$\begin{aligned}(x - (2 - i))(x - (2 + i)) &= x^2 - (2 + i)x - (2 - i)x + (2 - i)(2 + i) \\ &= x^2 - 4x + 5.\end{aligned}$$

Method 2: Sum and product of roots. The sum of the roots of $ax^2 + bx + c = 0$ is $-b/a$ and the product of the roots is c/a. Our work in Method 1 above shows why this result is true when $a = 1$. Specifically, we can write a monic quadratic with roots r and s as $(x - r)(x - s)$. Expanding this product gives $x^2 - (r + s)x + rs$, so the coefficient of the linear term is the opposite of the sum of the roots, and the constant term is the product of the roots. Since our roots in this case are $2 - i$ and $2 + i$, the sum is 4 and the product is 5, which means the desired quadratic is $x^2 - 4x + 5$.

Similarly, because $1 + 4i$ is a root of f, so is $1 - 4i$. The sum of these roots is 2 and their product is 17, so $1 + 4i$ and $1 - 4i$ are the roots of the quadratic $x^2 - 2x + 17$.

Since f must be divisible by both $x^2 - 4x + 5$ and $x^2 - 2x + 17$, and these two quadratics have no linear factors in common, the product of these two monic quadratics produces the desired monic polynomial f with smallest possible degree:

$$f(x) = (x^2 - 4x + 5)(x^2 - 2x + 17) = x^4 - 6x^3 + 30x^2 - 78x + 85.$$

□

Problem 6.28: Suppose $a + bi$ is a solution of the polynomial equation

$$c_4z^4 + ic_3z^3 + c_2z^2 + ic_1z + c_0 = 0,$$

where c_0, c_1, c_2, c_3, c_4, a, and b are real constants. Prove that $-a + bi$ is also a solution. *(Source: AHSME)*

Solution for Problem 6.28: Let $g(z) = c_4z^4 + ic_3z^3 + c_2z^2 + ic_1z + c_0$. The coefficients of $g(z)$ are not all real, so we can't deduce that the nonreal roots come in conjugate pairs. We don't know much about roots of polynomials with nonreal coefficients, so we'd like to relate $g(z)$ to a polynomial in which all the coefficients are real. If we had such a polynomial, we'd at least know that its nonreal roots come in conjugate pairs.

We notice that the imaginary coefficients are in the terms with odd powers of z, so if we let $z = xi$, the resulting coefficients of x are real:

$$g(xi) = c_4x^4 + c_3x^3 - c_2x^2 - c_1x + c_0.$$

Aha! Now, we've related $g(z)$ to a polynomial with real coefficients. We let

$$f(x) = g(xi) = c_4x^4 + c_3x^3 - c_2x^2 - c_1x + c_0$$

and investigate the roots of $f(x)$. We are given that $g(a + bi) = 0$. Because $f(x) = g(xi)$, we therefore know that $f(x)$ is 0 when $xi = a + bi$. Solving for x gives us $x = (a + bi)/i = b - ai$, and we have

$$f(b - ai) = g((b - ai)i) = g(a + bi) = 0.$$

Since $f(b - ai) = 0$ and $f(x)$ has real coefficients, we know that $\overline{b - ai} = b + ai$ is also a root of $f(x)$. Now, because $g(xi) = f(x) = 0$ when $x = b + ai$, we see that $g(z) = 0$ when $z = xi = (b + ai)i = -a + bi$, as desired. □

Exercises

6.4.1 Find all roots of $g(x) = 2x^3 + 5x^2 + 15x + 18$.

6.4.2 Find all roots of $f(t) = 2t^4 - 23t^2 + 27t - 36$.

6.4.3 Let $g(x)$ be a polynomial with real coefficients such that $g(2 + i) = g(3 + i) = 0$, and $\deg g \leq 4$. Other than $2 + i$ and $3 + i$, find all possible roots of g.

6.4.4 Prove that a fourth degree polynomial with a real root and real coefficients must have another real root (that may be the same as the first real root).

6.4.5⋆ A polynomial of degree four with leading coefficient 1 and integer coefficients has two real zeros, both of which are integers. Which of the following can also be a zero of the polynomial?

$$\frac{1+i\sqrt{11}}{2}, \frac{1+i}{2}, \frac{1}{2}+i, 1+\frac{i}{2}, \frac{1+i\sqrt{13}}{2}.$$

(Source: AMC 12) **Hints:** 2

6.4.6⋆ Find all values of x such that $x^4 + 5x^2 + 4x + 5 = 0$. **Hints:** 262

6.5 Summary

A **complex number** is a number of the form $a + bi$, where a and b are real numbers and $i^2 = -1$. We call a the **real part** of $a + bi$ and we call b the **imaginary part**. If $b \neq 0$, then $a + bi$ is **nonreal**. If $a = 0$, then the complex number $a + bi$ is called an **imaginary number**. Addition and multiplication of complex numbers are defined as follows:

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$
$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

Important: For any complex numbers v, w, and z, we have:

- $w + z = z + w$ and $wz = zw$
- $v + (w + z) = (v + w) + z$ and $(vw)z = v(wz)$
- $v(w + z) = vw + vz$

Important: The powers of i repeat in cycles of 4:

$$\begin{aligned} i^1 &= i^5 = i^9 = \cdots = i, \\ i^2 &= i^6 = i^{10} = \cdots = -1, \\ i^3 &= i^7 = i^{11} = \cdots = -i, \\ i^4 &= i^8 = i^{12} = \cdots = 1. \end{aligned}$$

If $z = a + bi$, then $a - bi$ is called the **conjugate** of z, denoted $\overline{z}$. So, we have $\overline{a + bi} = a - bi$. Together, a nonreal number and its conjugate are referred to as a **conjugate pair**. The product of a complex number and its conjugate is a real number. For any complex numbers z and w, we have $\overline{z + w} = \overline{z} + \overline{w}$ and $\overline{zw} = \overline{z} \cdot \overline{w}$.

Important: If w and z are complex numbers and $z \neq 0$, then we express the quotient w/z as a complex number by multiplying both the numerator and denominator by the conjugate of the denominator, $\overline{z}$:

$$\frac{w}{z} = \frac{w}{z} \cdot \frac{\overline{z}}{\overline{z}} = \frac{w\overline{z}}{z\overline{z}}.$$

Since $z\overline{z}$ is always real, we can express $w\overline{z}/(z\overline{z})$ as a complex number, which means we can write it in the form $a + bi$ for some pair of real numbers a and b.

A complex number can be represented as a point in the **complex plane**. Like the Cartesian plane, the complex plane has two axes: a horizontal **real axis** for the real part, and a vertical **imaginary axis** for the imaginary part. The real axis of the complex plane is labeled Re and the imaginary axis is labeled Im. Plotting the complex number $x + yi$ on the complex plane is the same as plotting the point (x, y) in the Cartesian plane.

Important: The **magnitude** of z, denoted by $|z|$, equals the distance from z to the origin on the complex plane. If $z = a + bi$, we have

$$|z| = \sqrt{a^2 + b^2} = \sqrt{z \cdot \bar{z}}.$$

The distance between complex numbers w and z on the complex plane is $|w - z|$.

Important: We often use $\mathrm{Re}(z)$ to refer to the real part of the complex number z, and we use $\mathrm{Im}(z)$ to refer to the imaginary part of z. So, for example $\mathrm{Re}(2 - 4i) = 2$ and $\mathrm{Im}(2 - 4i) = -4$. For any complex number z, we have

$$\mathrm{Re}(z) = \frac{z + \bar{z}}{2} \quad \text{and} \quad \mathrm{Im}(z) = \frac{z - \bar{z}}{2i}.$$

The **Fundamental Theorem of Algebra** states that every one-variable polynomial with complex coefficients has at least one complex root.

Important: Every polynomial of degree n has exactly n complex roots. (A number may appear multiple times among the roots of the polynomial. For example, the polynomial $(x - 1)(x - 2)^2$ has the three complex roots 1, 2, and 2.)

Important: Let $p(x)$ be a polynomial with real coefficients. If z is a root of $p(x)$, then $\bar{z}$ is also a root of $p(x)$. In other words, the nonreal roots of $p(x)$ come in conjugate pairs.

Things To Watch Out For!

WARNING!! We have to be careful about using notation like $\sqrt{-81}$. The real-valued function $f(x) = \sqrt{x}$ is only defined for $x \geq 0$. When we write $\sqrt{-81}$ as an intermediate step to solving an equation like $z^2 = -81$, we should keep in mind that we mean $i\sqrt{81}$.

Problem Solving Strategies

Concepts:

- If you are given a problem involving a general complex number z, it may help to write z in the form $a + bi$, and then express the problem in terms of a and b.
- It's sometimes easier to work with the square of the magnitude of a complex number than with the magnitude itself, in part because we can then apply the relationship $|z|^2 = z\bar{z}$.
- Trying specific examples can be a good way to discover general relationships.
- Many complex number concepts have a strong geometric significance as well as an algebraic one, so if you have trouble understanding an algebraic relationship regarding complex numbers, try considering it in the complex plane.

REVIEW PROBLEMS

6.29 Let $w = 2 + 3i$ and $z = 4 - 5i$. Express each of the following as a complex number:

(a) $2w - 3z$ (c) $\dfrac{2}{\overline{w} + z}$ (e) $\dfrac{-y + xi}{x + yi}$

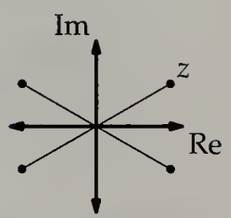
(b) $\dfrac{1}{w}$ (d) $\dfrac{w^3 + 2w^2z + wz^2}{w^2z + wz^2}$ (f) $\dfrac{(1-i)^4}{(1+i)^3}$

6.30 Solve the equation $-3 - x^2 = 9 + x^2$.

6.31 Express $\dfrac{1}{1 + \dfrac{1}{1 - \dfrac{1}{1+i}}}$ in the form $a + bi$, where a and b are real numbers.

6.32 Find the complex number z such that $\dfrac{z}{1+z} = -1 + i$.

6.33 Four complex numbers are plotted in the plane, as shown. One of them is z. The other three are $-z$, $\overline{z}$, and $-\overline{z}$. Which is which?



6.34 Let w and z be complex numbers.

(a) Show that $w\overline{z} + \overline{w}z$ is real.

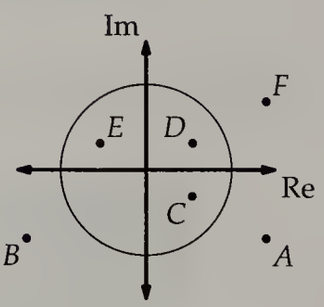
(b) Show that $w\overline{z} - \overline{w}z$ is imaginary.

6.35 Find two complex numbers whose squares equal $5 - 12i$.

6.36 Find all real numbers c such that $7 + i$ is 5 units from $10 + ci$ on the complex plane.

6.37 Find the area of the region enclosed by the graph of $|z - 4 + 5i| = 2\sqrt{3}$.

6.38 The diagram to the right shows several numbers in the complex plane. The circle has radius 1 and is centered at the origin. One of these numbers is the reciprocal of F. Which one? *(Source: AHSME)*



6.39 Find all complex numbers z such that $|z + 1 - i| = |z - 2|$.

6.40 The product of the complex numbers $a + bi$ and $c + di$ is real, where a, b, c, and d are real numbers. Prove that the product of the complex numbers $b + ai$ and $d + ci$ is also real.

6.41 Describe the graph of each of the following:

(a) $z - \overline{z} = -8i$ (b) $(4 - i)z - (4 + i)\overline{z} = 16i$ (c) $|7 + i - 2z| = 4$

6.42 Graph $|z - 5 + 2i| \leq 4$ on the complex plane.

6.43 Simplify $(i + 1)^{3200} - (i - 1)^{3200}$.

6.44 Show that $|z - 1|^2 + |z + 1|^2 = 4$ for all z such that $|z| = 1$.

6.45

(a) Prove that for any complex numbers w and z, we have $2(|w|^2 + |z|^2) = |w - z|^2 + |w + z|^2$.

(b) The relationship in part (a) is sometimes referred to as the **parallelogram law**. Why?

6.46 One of the roots of $x^4 + 9x^3 + 48x^2 + 78x - 136 = 0$ is $-3 + 5i$. Find the other three roots.

6.47 Let $f(x)$ be a polynomial with degree 6 and rational coefficients. The polynomial has five distinct roots: 1, $2 - i$, $2 + i$, $1 - i$, and $1 + i$.

(a) We call a value that appears twice among the roots of a polynomial a **double root** of the polynomial. Explain why f has a double root.

(b) Which root is the double root?

6.48 Let $r_1 = a + bi$ and $r_2 = a - bi$, where a and b are nonzero real numbers. Find the minimum degree of a polynomial with real coefficients and roots $r_1 i$ and $r_2 i$.

Challenge Problems

6.49 Evaluate $i + 2i^2 + 3i^3 + 4i^4 + \cdots + 64i^{64}$.

6.50 Let a, b, and c be complex numbers. Suppose we want to solve the equation $az + b\bar{z} = c$ for z.

(a) Show that if $|a| \neq |b|$, then the equation has a unique solution z.

(b) Describe the solutions when $a = -1 + 3i$, $b = 1 + 3i$, and $c = i$.

6.51 Find all complex numbers z such that $|z - 1| = |z + 3| = |z - i|$. *(Source: ARML)* **Hints:** 121

6.52 Find the area of the region of points z in the complex plane such that

$$|z + 4 - 4i| \leq 4\sqrt{2} \qquad \text{and} \qquad |z - 4 - 4i| \geq 4\sqrt{2}.$$

6.53 Find all roots of the polynomial $iy^3 - 8y^2 - 22iy + 21$.

6.54 Show that the points w, x, y, and z are the vertices of a parallelogram in the complex plane if and only if the sum of some two of them is equal to the sum of the other two. **Hints:** 164

6.55★ Let u and v be complex numbers such that $|u| = |v| = 1$, $u \neq v$, and $u \neq -v$. Show that $\dfrac{4uv}{(u+v)^2}$ is real, and that $\dfrac{u+v}{u-v}$ is imaginary. **Hints:** 102

6.56★ Let (x, y) be a pair of real numbers satisfying

$$56x + 33y = -\frac{y}{x^2 + y^2} \quad \text{and} \quad 33x - 56y = \frac{x}{x^2 + y^2}.$$

Determine the value of $|x| + |y|$. **Hints:** 94

6.57★ Suppose we define a sequence of complex numbers by $z_1 = 0$ and $z_{n+1} = z_n^2 + i$ for $n \geq 1$. How far away from the origin is z_{111}? *(Source: AHSME)* **Hints:** 206

6.58★ Show that if $w_1 + w_2 + w_3 = 0$ and $|w_1| = |w_2| = |w_3| = 1$, then $w_1^2 + w_2^2 + w_3^2 = 0$. **Hints:** 11, 69, 60

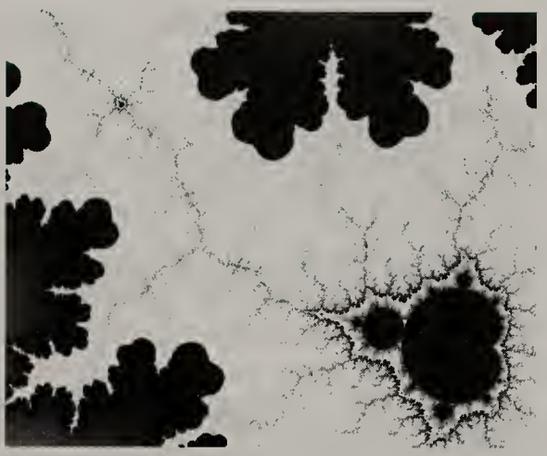
6.59★ A function f is defined on the complex numbers by $f(z) = (a + bi)z$, where a and b are positive numbers. This function has the property that the image of each point in the complex plane is equidistant from that point and the origin. Given that $|a + bi| = 8$, find the value of b^2. *(Source: AIME)* **Hints:** 143, 125

6.60 For all nonzero complex numbers z, let $f(z) = 1/\bar{z}$.

(a) Show that $f(f(z)) = z$ for all $z \neq 0$.

(b)★ As z varies along the line $(1 + 2i)z - (1 - 2i)\bar{z} = i$, what curve does $f(z)$ trace? **Hints:** 268

6.61★ Find all ordered pairs (a, b) such that $a + \dfrac{a + 8b}{a^2 + b^2} = 2$ and $b + \dfrac{8a - b}{a^2 + b^2} = 0$. **Hints:** 226



A complex system that works is invariably found to have evolved from a simple system that worked. – John Gall

CHAPTER 7

Trigonometry and Complex Numbers

7.1 Polar Form of Complex Numbers

In Chapter 5, we learned how to convert rectangular coordinates (x, y) into polar coordinates (r, θ). In this section, we apply this conversion technique to complex numbers to write complex numbers in **polar form**.

Problems

Problem 7.1: Explain why every complex number can be written in the form $r(\cos\theta + i\sin\theta)$ for some constants r and θ with $r \geq 0$. We call this form the **polar form** of a complex number.

Problem 7.2: Suppose $z = r(\cos\theta + i\sin\theta)$, where $r \geq 0$. What is $|z|$ in terms of r and θ?

Problem 7.3: Write each of the following in polar form by finding the positive value of r and the value of θ with $0 \leq \theta < 2\pi$ such that the given complex number equals $r(\cos\theta + i\sin\theta)$.

(a) -6

(b) $4 + 4i$

(c) $3 - i\sqrt{3}$

(d) $-\sqrt{2} + i\sqrt{6}$

Problem 7.4: Let $w = 3 + i\sqrt{3}$ and $z = 2\sqrt{3} - 6i$.

(a) Find wz.

(b) Express w, z, and wz in polar form. Notice anything interesting?

Problem 7.5: Let $w = r(\cos\alpha + i\sin\alpha)$ and $z = s(\cos\beta + i\sin\beta)$. Express wz in polar form.

Problem 7.6:
(a) Prove that $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$ for any positive integer n and any angle θ.
(b) Prove that $(\cos\theta - i\sin\theta)^n = \cos n\theta - i\sin n\theta$ for any positive integer n and any angle θ.

Problem 7.7: Express $\cos 5\theta$ in terms of $\cos\theta$.

Problem 7.1: Explain why every complex number can be written in the form $r(\cos\theta + i\sin\theta)$ for some constants r and θ with $r \geq 0$. We call this form the **polar form** of a complex number.

Solution for Problem 7.1: Let z be the complex number $x + yi$, where x and y are real numbers. In Section 5.2, we learned that any point in the Cartesian plane with rectangular coordinates (x, y) can also be denoted with the polar coordinates (r, θ), where $x = r\cos\theta$ and $y = r\sin\theta$, with $r \geq 0$ and $0 \leq \theta < 2\pi$. Therefore, for any ordered pair of real numbers (x, y), there is a constant r and an angle θ such that $x = r\cos\theta$ and $y = r\sin\theta$. This means that we can write the complex number $x + yi$ in the form $r(\cos\theta + i\sin\theta)$ for some constants r and θ such that $r \geq 0$ and $0 \leq \theta < 2\pi$. □

The expression $\cos\theta + i\sin\theta$ is sometimes shortened to $\text{cis}\,\theta$. We'll use this notation sparingly in this text, because in the next section we will introduce a better way to express complex numbers in terms of r and θ.

Because the complex number $z = x+yi$ in the complex plane corresponds to rectangular coordinates (x, y) in the Cartesian plane, we call $x+yi$ the **rectangular form** of the complex number z. Similarly, because $z = r(\cos\theta+i\sin\theta)$ in the complex plane corresponds to the point in the Cartesian plane with polar coordinates (r, θ), we call this the **polar form** of z. We call r the **magnitude** of the complex number and θ the number's **argument**.

Don't we already have something called the magnitude of a complex number?

Problem 7.2: Suppose $z = r(\cos\theta + i\sin\theta)$, where $r \geq 0$. What is $|z|$ in terms of r and θ?

Solution for Problem 7.2: Solution 1: Algebra. We have

$$|z| = |r\cos\theta + ir\sin\theta| = \sqrt{(r\cos\theta)^2 + (r\sin\theta)^2} = \sqrt{r^2(\cos^2\theta + \sin^2\theta)} = \sqrt{r^2} = r.$$

(The step $\sqrt{r^2} = r$ is only valid because $r \geq 0$.)

Solution 2: Geometry. As we saw in Problem 7.1, the point $r(\cos\theta + i\sin\theta)$ in the complex plane corresponds to the point (r, θ) in the Cartesian plane. Since $|z|$ equals the distance from z to the origin in the complex plane, and the distance from (r, θ) to the origin in the Cartesian plane is r (when $r \geq 0$), we have $|z| = r$. □

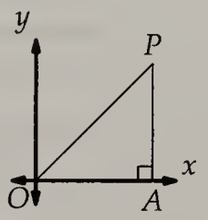
Problem 7.3: Write each of the following in polar form by finding the positive value of r and the value of θ with $0 \leq \theta < 2\pi$ such that the given complex number equals $r(\cos\theta + i\sin\theta)$.
(a) -6
(b) $4 + 4i$
(c) $3 - i\sqrt{3}$
(d) $-\sqrt{2} + i\sqrt{6}$

Solution for Problem 7.3:

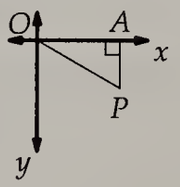
Important: Converting complex numbers in rectangular form to polar form is exactly the same as converting the rectangular coordinates (x, y) to polar coordinates (r, θ).

(a) The point -6 in the complex plane is 6 units to the left of the origin. In the Cartesian plane, the polar coordinates of the point 6 units to the left of the origin are $(6, \pi)$. Therefore, a polar form of the complex number -6 is $6(\cos\pi + i\sin\pi)$, or $6\,\text{cis}\,\pi$.

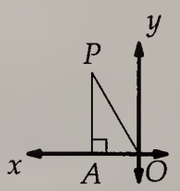
(b) The point $4+4i$ in the complex plane is point P in the diagram at right. Since $PA = AO = 4$, we know that $\triangle PAO$ is an isosceles right triangle. Therefore, we have $PO = 4\sqrt{2}$ and $\angle POA = \frac{\pi}{4}$, which means that a polar form of $4 + 4i$ is $4\sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)$, or $4\sqrt{2}\,\text{cis}\,\frac{\pi}{4}$.



(c) The point $3 - i\sqrt{3}$ in the complex plane is point P in the diagram at right. We have $PO = \sqrt{PA^2 + AO^2} = 2\sqrt{3}$, so the hypotenuse of $\triangle PAO$ is double the smaller leg of $\triangle PAO$. Therefore, we have $\angle POA = \frac{\pi}{6}$. Since P is $\frac{\pi}{6}$ clockwise from the positive x-axis, it is $\frac{11\pi}{6}$ radians counterclockwise from the x-axis. So, we can write $3 - i\sqrt{3}$ in polar form as $2\sqrt{3}\left(\cos\frac{11\pi}{6} + i\sin\frac{11\pi}{6}\right)$, or $2\sqrt{3}\,\text{cis}\,\frac{11\pi}{6}$.



(d) The point $-\sqrt{2} + i\sqrt{6}$ in the complex plane is point P in the diagram at right. We have $PO = \sqrt{PA^2 + AO^2} = 2\sqrt{2}$, so the hypotenuse of $\triangle PAO$ is double the smaller leg of $\triangle PAO$. Therefore we have $\angle POA = \frac{\pi}{3}$, which means that P is $\frac{2\pi}{3}$ radians counterclockwise from the positive x-axis. So, the polar form of $-\sqrt{2} + i\sqrt{6}$ is $2\sqrt{2}\left(\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}\right)$, or $2\sqrt{2}\,\text{cis}\,\frac{2\pi}{3}$.



□

Sidenote: You may have noticed that while we usually write complex numbers in the form $a + bi$, with the i after the imaginary part of the number, in the problem above we have written $3 - i\sqrt{3}$ and $-\sqrt{2} + i\sqrt{6}$ instead of $3 - \sqrt{3}i$ and $-\sqrt{2} + \sqrt{6}i$. We follow this convention in part to avoid confusion, since we might mistake the i in $3 - \sqrt{3}i$ as being inside the square root symbol. (This mistake is particularly easy to make when the numbers in question are hand-written.)

Problem 7.4: Let $w = 3 + i\sqrt{3}$ and $z = 2\sqrt{3} - 6i$.

(a) Find wz.

(b) Express w, z, and wz in polar form. Notice anything interesting?

Solution for Problem 7.4:

(a) We have

$$wz = (3 + i\sqrt{3})(2\sqrt{3} - 6i) = 6\sqrt{3} - 18i + 6i + 6\sqrt{3} = 12\sqrt{3} - 12i.$$

(b) We could take a geometric approach, as we did in Problem 7.3. Or, we can note by the forms of w, z, and wz that the arguments of these complex numbers are multiples of $\frac{\pi}{6}$. We can determine the corresponding

arguments by first factoring out the magnitudes of each, and then using our knowledge of trigonometry to find the arguments. Since $|w| = |3 + i\sqrt{3}| = 2\sqrt{3}$, we factor $2\sqrt{3}$ out from w to find

$$w = 3 + i\sqrt{3} = 2\sqrt{3}\left(\frac{\sqrt{3}}{2} + \frac{1}{2}i\right).$$

The argument of w then is the angle θ such that $\cos\theta = \frac{\sqrt{3}}{2}$ and $\sin\theta = \frac{1}{2}$, so we have $\theta = \frac{\pi}{6}$ and $w = 2\sqrt{3}\left(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}\right)$.

Similarly, $|z| = |2\sqrt{3} - 6i| = \sqrt{12 + 36} = 4\sqrt{3}$, and we have $z = 2\sqrt{3} - 6i = 4\sqrt{3}\left(\frac{1}{2} - \frac{\sqrt{3}}{2}i\right)$. This means the argument of z is the angle θ such that $\cos\theta = \frac{1}{2}$ and $\sin\theta = -\frac{\sqrt{3}}{2}$, which gives us $\theta = \frac{5\pi}{3}$. So, we have $z = 4\sqrt{3}\left(\cos\frac{5\pi}{3} + i\sin\frac{5\pi}{3}\right)$.

Finally, we have $|wz| = |12\sqrt{3} - 12i| = \sqrt{432 + 144} = 24$, and $wz = 24\left(\frac{\sqrt{3}}{2} - \frac{1}{2}i\right)$. Therefore, we seek θ such that $\cos\theta = \frac{\sqrt{3}}{2}$ and $\sin\theta = -\frac{1}{2}$, which gives us $\theta = \frac{11\pi}{6}$ and $wz = 24\left(\cos\frac{11\pi}{6} + i\sin\frac{11\pi}{6}\right)$.

Listing all of our results, we have

$$w = 2\sqrt{3}\left(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}\right),$$
$$z = 4\sqrt{3}\left(\cos\frac{5\pi}{3} + i\sin\frac{5\pi}{3}\right),$$
$$wz = 24\left(\cos\frac{11\pi}{6} + i\sin\frac{11\pi}{6}\right).$$

Interesting—the magnitude of wz is the product of the magnitudes of w and z, and the argument of wz is the sum of the arguments of w and z.

□

Is our observation at the end of this solution true for any two complex numbers w and z?

Problem 7.5: Let $w = r(\cos\alpha + i\sin\alpha)$ and $z = s(\cos\beta + i\sin\beta)$. Express wz in polar form.

Solution for Problem 7.5: We have

$$\begin{aligned} wz &= r(\cos\alpha + i\sin\alpha)\cdot s(\cos\beta + i\sin\beta) \\ &= rs\cdot(\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) \\ &= rs\cdot(\cos\alpha\cos\beta - \sin\alpha\sin\beta + i\sin\alpha\cos\beta + i\cos\alpha\sin\beta). \end{aligned}$$

We recognize the real part of the expression in parentheses as $\cos(\alpha + \beta)$ and the imaginary part as $\sin(\alpha + \beta)$, so we have

$$wz = rs\left(\cos(\alpha + \beta) + i\sin(\alpha + \beta)\right).$$

□

Important: For any two complex numbers w and z, the magnitude of wz is the product of the magnitudes of w and z, and the argument of wz is the sum of the arguments of w and z (or a multiple of 2π less than the sum of the arguments of w and z, if we restrict arguments to being in the interval $[0, 2\pi)$).

Problem 7.6:
(a) Prove that $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$ for any positive integer n and any angle θ.
(b) Prove that $(\cos\theta - i\sin\theta)^n = \cos n\theta - i\sin n\theta$ for any positive integer n and any angle θ.

Solution for Problem 7.6:

(a) We have already done the heavy lifting for this problem when we solved Problem 7.5. There, we showed that $(\text{cis}\,\alpha)(\text{cis}\,\beta) = \text{cis}(\alpha+\beta)$. So, we have

$$(\cos\theta + i\sin\theta)^2 = (\cos\theta + i\sin\theta)(\cos\theta + i\sin\theta) = \cos(\theta+\theta) + i\sin(\theta+\theta) = \cos 2\theta + i\sin 2\theta.$$

We multiply this result by $\cos\theta + i\sin\theta$ to find an expression for $(\cos\theta + i\sin\theta)^3$:

$$\begin{aligned}(\cos\theta + i\sin\theta)^3 &= (\cos\theta + i\sin\theta)(\cos\theta + i\sin\theta)^2\\ &= (\cos\theta + i\sin\theta)(\cos 2\theta + i\sin 2\theta)\\ &= \cos(\theta+2\theta) + i\sin(\theta+2\theta)\\ &= \cos 3\theta + i\sin 3\theta.\end{aligned}$$

We use induction to extend this argument to any integer n. Clearly, we have $(\cos\theta+i\sin\theta)^n = \cos n\theta+i\sin n\theta$ for $n = 1$. Assuming that $(\cos\theta + i\sin\theta)^k = \cos k\theta + i\sin k\theta$ for some positive integer k, we have

$$\begin{aligned}(\cos\theta + i\sin\theta)^{k+1} &= (\cos\theta + i\sin\theta)(\cos\theta + i\sin\theta)^k\\ &= (\cos\theta + i\sin\theta)(\cos k\theta + i\sin k\theta)\\ &= \cos(\theta + k\theta) + i\sin(\theta + k\theta)\\ &= \cos((k+1)\theta) + i\sin((k+1)\theta).\end{aligned}$$

Therefore, by induction, we have $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$ for all positive integers n.

(b) The complex numbers in this part are not in polar form, but they are the conjugates of the complex numbers in part (a). This observation gives us a quick solution. If two complex numbers are equal, their conjugates must be equal. So, taking the conjugate of both sides of

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$$

gives

$$\overline{(\cos\theta + i\sin\theta)^n} = \overline{\cos n\theta + i\sin n\theta} = \cos n\theta - i\sin n\theta.$$

Since $\overline{z^n} = (\overline{z})^n$, we have

$$\overline{(\cos\theta + i\sin\theta)^n} = (\overline{\cos\theta + i\sin\theta})^n = (\cos\theta - i\sin\theta)^n,$$

so we have the desired $(\cos\theta - i\sin\theta)^n = \cos n\theta - i\sin n\theta$.

□

The relationship we proved in part (a) is commonly referred to as **de Moivre's Theorem**.

Important: **De Moivre's Theorem** states that

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$$

for any θ and any positive integer n.

This relationship is historically important because it helped pave the way for a deeper understanding of complex numbers that we explore in the next section. Because de Moivre's Theorem is a straightforward result of the concepts we'll develop in the next section, you don't need to memorize it now.

Problem 7.7: Express $\cos 5\theta$ in terms of $\cos\theta$.

Solution for Problem 7.7: We could start by writing $\cos 5\theta = \cos(4\theta + \theta)$ and using the angle sum identity, but that would give us an expression that includes $\cos 4\theta$ and $\sin 4\theta$. We could then write 4θ as $3\theta + \theta$ and apply angle sum identities again, but that will create quite a mess. Let's see if we can find a better approach.

As a starting point, we seek an equation that involves both $\cos 5\theta$ and $\cos\theta$. De Moivre's Theorem gives us one:

$$\cos 5\theta + i\sin 5\theta = (\cos\theta + i\sin\theta)^5.$$

Now, we just have to expand the right side. We can use the Binomial Theorem or a bit of algebra to see that

$$(x+y)^5 = x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5.$$

We expand $(\cos\theta + i\sin\theta)^5$ by letting $x = \cos\theta$ and $y = i\sin\theta$ in our expansion of $(x+y)^5$. This gives us

$$\cos 5\theta + i\sin 5\theta = \cos^5\theta + 5i\cos^4\theta\sin\theta - 10\cos^3\theta\sin^2\theta - 10i\cos^2\theta\sin^3\theta + 5\cos\theta\sin^4\theta + i\sin^5\theta.$$

The real part of the left must equal the real part of the right. Combining this with the fact that $\sin^2\theta = 1 - \cos^2\theta$ gives us

$$\begin{aligned}
\cos 5\theta &= \cos^5\theta - 10\cos^3\theta\sin^2\theta + 5\cos\theta\sin^4\theta \\
&= \cos^5\theta - 10\cos^3\theta(1-\cos^2\theta) + 5\cos\theta(1-\cos^2\theta)^2 \\
&= \cos^5\theta - 10\cos^3\theta + 10\cos^5\theta + 5\cos\theta(1 - 2\cos^2\theta + \cos^4\theta) \\
&= \cos^5\theta - 10\cos^3\theta + 10\cos^5\theta + 5\cos\theta - 10\cos^3\theta + 5\cos^5\theta \\
&= 16\cos^5\theta - 20\cos^3\theta + 5\cos\theta.
\end{aligned}$$

□

Concept: De Moivre's Theorem can help us tackle problems involving $\cos n\theta$ or $\sin n\theta$, where n is a positive integer greater than 1.

Sidenote: Above, we showed that

$$\cos 5\theta = 16\cos^5\theta - 20\cos^3\theta + 5\cos\theta.$$

So, if we let $f(x) = 16x^5 - 20x^3 + 5x$, then $\cos 5\theta = f(\cos\theta)$. This polynomial $f(x)$ is one of a family of polynomials called the **Chebyshev polynomials of the first kind**. These polynomials are typically denoted T_n, and satisfy the definition $T_0(x) = 1$, $T_1(x) = x$, and

$$T_{n+1} = 2xT_n(x) - T_{n-1}(x).$$

You'll have a chance to explore Chebyshev polynomials in Challenge Problem 7.69, where you'll show that $T_n(\cos\theta) = \cos n\theta$ for any positive integer n and any angle θ. (In some sources, "Chebyshev" is given the alternative spelling "Tchebychev," which may explain why we typically use "T" for the Chebyshev polynomials.)

Exercises

7.1.1 Write each of the following in polar form:

(a) $-32i$

(b) $2 - 2i$

(c) $2\sqrt{3} + 6i$

(d) $\frac{1}{4} - \frac{\sqrt{3}}{4}i$

7.1.2 Write each of the following in rectangular form:

(a) $6 \operatorname{cis} 150°$

(b) $8 \operatorname{cis} 3780°$

(c) $2 \operatorname{cis} \frac{2\pi}{3}$

(d) $9 \operatorname{cis} \frac{3\pi}{4}$

7.1.3 Show that if $w = r \operatorname{cis} \alpha$ and $z = s \operatorname{cis} \beta$ (and $z \neq 0$), then $\frac{w}{z} = \frac{r}{s} \operatorname{cis}(\alpha - \beta)$.

7.1.4 Evaluate each of the following in your head without using paper, pencil, pen, or a calculator.

(a) $(6\sqrt{3} + 18i)(-2\sqrt{3} + 6i)$

(b) $(4 + 4i)(-3 - 3i)(-10 + 10i)$

7.1.5 For what values of θ does $\operatorname{cis} \theta = \operatorname{cis}(-\theta)$?

7.1.6 Test the identity we found in Problem 7.7 for $\cos 5\theta$ by letting $\theta = 30°$.

7.1.7 Express $\cos 7\theta$ in terms of $\cos \theta$.

7.1.8★ For how many positive integers n less than or equal to 1000 is

$$(\sin t + i \cos t)^n = \sin nt + i \cos nt$$

true for all real t? *(Source: AIME)* **Hints:** 52

7.2 Exponential Form of Complex Numbers

In the last section, we saw that the argument of a product of complex numbers is the sum of the arguments of the complex numbers. This observation inspires another notation for complex numbers.

Problems

Problem 7.8: Let $f(x) = \cos x + i \sin x$.

(a) Show that $f(0) = 1$ and $f(x)f(y) = f(x + y)$ for all real x and y.

(b) Suppose a function g satisfies $g(0) = 1$ and $g(x)g(y) = g(x+y)$ for all real x and y. Show that it also satisfies the following:

(i) $g(x) \neq 0$ for all x.

(ii) $g(x - y) = g(x)/g(y)$ for all real numbers x and y.

(iii) $[g(x)]^n = g(nx)$ for any real number x and any positive integer n.

(iv) $g(-x) = 1/g(x)$.

(c) What other functions g have $g(0) = 1$ and satisfy the identity $g(x)g(y) = g(x + y)$ for all x and y?

Problem 7.8: Let $f(x) = \cos x + i \sin x$.

(a) Show that $f(0) = 1$ and $f(x)f(y) = f(x + y)$ for all real x and y.

(b) Suppose a function g satisfies $g(0) = 1$ and $g(x)g(y) = g(x + y)$ for all real x and y. Show that it also satisfies the following:

(i) $g(x) \neq 0$ for all x.

(ii) $g(x - y) = g(x)/g(y)$ for all real numbers x and y.

(iii) $[g(x)]^n = g(nx)$ for any real number x and any positive integer n.

(iv) $g(-x) = 1/g(x)$.

(c) What other functions g have $g(0) = 1$ and satisfy the identity $g(x)g(y) = g(x + y)$ for all x and y?

Solution for Problem 7.8:

(a) We have $f(0) = \cos 0 + i \sin 0 = 1 + 0i = 1$. Proving that $f(x)f(y) = f(x + y)$ is essentially a simplified version of Problem 7.5:

$$\begin{aligned} f(x)f(y) &= (\cos x + i \sin x)(\cos y + i \sin y) \\ &= \cos x \cos y - \sin x \sin y + i(\sin x \cos y + \cos x \sin y) \\ &= \cos(x + y) + i \sin(x + y) \\ &= f(x + y). \end{aligned}$$

(b) Letting $y = -x$ in $g(x)g(y) = g(x + y)$ gives $g(x)g(-x) = g(0) = 1$ for all x. Therefore, we cannot have $g(x) = 0$ for any x, since $g(x)g(-x)$ cannot be 0.

Since $g(x)g(y) = g(x + y)$, dividing both sides by $g(y)$ gives $g(x) = g(x + y)/g(y)$. Letting $x + y = z$, so that $x = z - y$, gives us $g(z - y) = g(z)/g(y)$, which proves (ii).

The identity $[g(x)]^n = g(nx)$ is simply a repeated application of $g(x)g(y) = g(x + y)$. Specifically, we have $[(g(x)]^2 = g(x)g(x) = g(2x)$, and $[g(x)]^3 = [g(x)]^2 g(x) = g(2x)g(x) = g(2x + x) = g(3x)$, and so on. We can formalize this with induction. We clearly have $[g(x)]^1 = g(x)$, and if $[g(x)]^k = g(kx)$ for some positive integer k, then we have

$$[g(x)]^{k+1} = g(x)[g(x)]^k = g(x)g(kx) = g(x + kx) = g((k + 1)x).$$

This shows that if $[g(x)]^n = g(nx)$ holds for $n = k$, then it holds for $n = k + 1$. We have already showed that it holds for $n = 1$, so by induction, it holds for all positive integers n. (Perhaps you recognize this argument from our proof of de Moivre's Theorem.)

Since $g(0) = 1$ and $g(x)/g(y) = g(x - y)$, letting $x = 0$ gives us $1/g(y) = g(-y)$, which proves identity (iv).

(c) Exponential functions have this behavior! For example, if $g(x) = 2^x$, we have $g(0) = 2^0 = 1$ and

$$g(x)g(y) = 2^x \cdot 2^y = 2^{x+y} = g(x + y).$$

Notice that exponential functions also satisfy the identities described in part (b); these are examples of the laws of exponents with which you're already familiar.

□

We see that the arguments of complex numbers appear to behave a lot like exponents. Specifically, we showed in Problem 7.8 that the function

$$f(x) = \cos x + i \sin x$$

satisfies the equation

$$f(x)f(y) = f(x + y)$$

for all x and y. Moreover, we confirmed that $f(0) = 1$, and that f satisfies many of the relationships we know as the laws of exponents, such as $f(x)/f(y) = f(x-y)$, $f(-x) = 1/f(x)$, and $[f(x)]^n = f(nx)$ for all positive integers n.

Because $f(x) = \cos x + i\sin x$ satisfies the same relationships that exponential functions satisfy, we have a special exponential notation for complex numbers. We write $\cos x + i\sin x$ as e^{ix}.

Definition: For all real x, we define e^{ix} as

$$e^{ix} = \cos x + i\sin x.$$

With this notation, we can write the identities we showed in Problem 7.8 as

$$e^{ix} \cdot e^{iy} = e^{i(x+y)},$$
$$e^{ix}/e^{iy} = e^{i(x-y)},$$
$$\left(e^{ix}\right)^n = e^{inx},$$
$$e^{i(-x)} = \frac{1}{e^{ix}}.$$

Now you see why we use this notation—all of these relationships are easy to recall because they are the same as the laws of exponents. For example, if you see $e^{i\alpha}/e^{i\beta}$ in a problem, it's easy to recall that this equals $e^{i(\alpha-\beta)}$, but if you see $(\cos\alpha + i\sin\alpha)/(\cos\beta + i\sin\beta)$ in a problem, it may not be immediately obvious that this expression equals $\cos(\alpha-\beta) + i\sin(\alpha-\beta)$. You might also have noticed that for positive integers n, the relationship $\left(e^{ix}\right)^n = e^{inx}$ is de Moivre's Theorem. This is why we suggested not memorizing de Moivre's Theorem after we proved it on page 223. Exponential form makes de Moivre's Theorem easy to remember.

Because $\cos x + i\sin x = e^{ix}$, we can write the polar form of a complex number $z = r(\cos\theta + i\sin\theta)$ in the form

$$z = re^{i\theta}.$$

We call this the **exponential form** of the complex number z, where r is the magnitude of z and θ is the argument, as with polar form.

WARNING!! We always use radians in exponential notation, never degrees.

You'll have to wait until your study of calculus to see technical reasons for using radians rather than degrees, but our introduction of radians in Section 2.2 gives a qualitative explanation why we prefer radians.

Sidenote: The symbol e is not just a placeholder to remind us that complex numbers can be written in an exponential form. There is a real number e such that we have $e^{i\theta} = \cos\theta + i\sin\theta$. This real number is irrational and approximately equal to 2.7182818.

There are many ways to define e. One of the simplest is to consider the expression $g(n) = \left(1 + \frac{1}{n}\right)^n$. Using a calculator or computer, compute $g(n)$ for larger and larger values of n. As n gets *very* large, you'll find that $g(n)$ appears to become constant. That constant is e.

While you'll find e very useful in your study of complex numbers, you'll see many more wonderful properties of e in your continued study of mathematics, and you'll learn why it is considered one of the most important constants in mathematics.

Problems

Problem 7.9: Express each of the following in exponential form:

(a) $-2i$ (b) $1-i$ (c) $-3\sqrt{3}+3i$

Problem 7.10: Express each of the following in rectangular form:

(a) $e^{\pi i}$ (b) $6e^{\pi i/3}$ (c) $4e^{7\pi i/4}$

Problem 7.11: Evaluate the following:

(a) $(1+i)^{20}$ (b) $(3-i\sqrt{3})^5$

Problem 7.12: For what values of θ is $e^{i\theta}=1$?

Problem 7.13: Show that if $z=re^{i\theta}$, then $\overline{z}=re^{-i\theta}$.

Problem 7.14:

(a) Use exponential form to prove that $z\overline{z}=|z|^2$ for all complex numbers z.

(b) Use exponential form to prove that $|zw|=|z||w|$ for all complex numbers z and w.

(c) Prove that complex number multiplication is associative. In other words, show that if w_1, w_2, and w_3 are complex numbers, then $(w_1w_2)w_3=w_1(w_2w_3)$.

Problem 7.15: Express $e^{ix}\cdot e^{iy}$ in terms of $\cos x$, $\sin x$, $\cos y$, and $\sin y$. How does your result make remembering the angle sum identities for $\cos(x+y)$ and $\sin(x+y)$ easy?

Problem 7.16:

(a) Find $\cos\theta$ in terms of $e^{i\theta}$ and $e^{-i\theta}$.

(b) Explain how you can use your relationship from part (a) to remember the product-to-sum identity for $\cos x\cos y$.

Problem 7.17: Find the positive integer n such that

$$\arctan\frac{1}{3}+\arctan\frac{1}{4}+\arctan\frac{1}{5}+\arctan\frac{1}{n}=\frac{\pi}{4}.$$

(Source: AIME) **Hints:** 74, 124

Problem 7.9: Express each of the following in exponential form:

(a) $-2i$ (b) $1-i$ (c) $-3\sqrt{3}+3i$

Solution for Problem 7.9: Converting rectangular form to exponential form is essentially the same as converting rectangular form to polar form.

(a) We have $|-2i|=2$. Since $-2i$ is 2 units below the origin in the complex plane, the argument of $-2i$ in exponential (and polar) form is $\frac{3\pi}{2}$. Therefore, we have $-2i=2e^{3\pi i/2}$.

(b) We have $|1 - i| = \sqrt{2}$, so we convert $1 - i$ to exponential form by writing $1 - i = \sqrt{2}\left(\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}i\right)$ and finding the angle θ such that $\cos\theta = \frac{\sqrt{2}}{2}$ and $\sin\theta = -\frac{\sqrt{2}}{2}$. This angle is $\frac{7\pi}{4}$, so we have $1 - i = \sqrt{2}e^{7\pi i/4}$.

(c) We have $|-3\sqrt{3} + 3i| = 6$, so we convert $-3\sqrt{3} + 3i$ to exponential form by writing $-3\sqrt{3} + 3i = 6\left(-\frac{\sqrt{3}}{2} + \frac{i}{2}\right)$ and finding the angle θ such that $\cos\theta = -\frac{\sqrt{3}}{2}$ and $\sin\theta = \frac{1}{2}$. This angle is $\frac{5\pi}{6}$, so we have $-3\sqrt{3} + 3i = 6e^{5\pi i/6}$.

□

Problem 7.10: Express each of the following in rectangular form:

(a) $e^{\pi i}$ (b) $6e^{\pi i/3}$ (c) $4e^{7\pi i/4}$

Solution for Problem 7.10:

(a) We have $e^{\pi i} = \cos\pi + i\sin\pi = -1 + 0i = -1$.

Sidenote: Adding 1 to both sides of $e^{\pi i} = -1$ gives us an equation that relates the five most important constants in mathematics:

$$e^{\pi i} + 1 = 0.$$

(b) We have $6e^{\pi i/3} = 6\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right) = 6\left(\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) = 3 + 3i\sqrt{3}$.

(c) We have $4e^{7\pi i/4} = 4\left(\cos\frac{7\pi}{4} + i\sin\frac{7\pi}{4}\right) = 4\left(\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}i\right) = 2\sqrt{2} - 2i\sqrt{2}$.

□

Problem 7.11: Evaluate the following:

(a) $(1 + i)^{20}$ (b) $(3 - i\sqrt{3})^5$

Solution for Problem 7.11:

(a) We could start with the Binomial Theorem to expand $(1+i)^{20}$ . . . on second thought, maybe not. Exponential form offers a much easier way. In exponential form, $1 + i$ is $\sqrt{2}e^{\pi i/4}$, so we have

$$(1 + i)^{20} = (\sqrt{2}e^{\pi i/4})^{20} = (2^{1/2})^{20}e^{20\pi i/4} = 2^{10}e^{5\pi i} = 1024(\cos 5\pi + i\sin 5\pi) = -1024.$$

(b) As in the previous part, we convert $3 - i\sqrt{3}$ to exponential form, writing it as $2\sqrt{3}e^{11\pi i/6}$. We therefore have

$$(3 - i\sqrt{3})^5 = (2\sqrt{3}e^{11\pi i/6})^5 = (2\sqrt{3})^5 e^{55\pi i/6} = 32 \cdot 9\sqrt{3}\left(\cos\frac{55\pi}{6} + i\sin\frac{55\pi}{6}\right).$$

Subtracting multiples of 2π repeatedly from $\frac{55\pi}{6}$ gives us

$$(3 - i\sqrt{3})^5 = 288\sqrt{3}\left(\cos\frac{7\pi}{6} + i\sin\frac{7\pi}{6}\right) = 288\sqrt{3}\left(-\frac{\sqrt{3}}{2} - \frac{i}{2}\right) = -432 - 144i\sqrt{3}.$$

□

Problem 7.12: For what values of θ is $e^{i\theta} = 1$?

Solution for Problem 7.12: We obviously have $e^{i\theta} = 1$ when $\theta = 0$, but are there other values of θ for which $e^{i\theta} = 1$? Writing $e^{i\theta}$ in polar form gives us $\cos\theta + i\sin\theta = 1$, which means that $\cos\theta = 1$ and $\sin\theta = 0$. Therefore, θ must be an integer multiple of 2π in order to have $e^{i\theta} = 1$.

We also could have reasoned to this answer by thinking about the complex plane. Plotting the point $e^{i\theta}$ in the complex plane is the same as plotting the point with polar coordinates $(1, \theta)$ in the Cartesian plane. So, the point $e^{i\theta}$ is on the unit circle graphed in the complex plane. Specifically, it is the terminal point of θ, which means we have $e^{i\theta} = 1$ if and only if θ is an integer multiple of 2π. □

Problem 7.13: Show that if $z = re^{i\theta}$, then $\overline{z} = re^{-i\theta}$.

Solution for Problem 7.13: We have $z = re^{i\theta} = r(\cos\theta + i\sin\theta)$, so

$$\overline{z} = \overline{r\cos\theta + ri\sin\theta} = r\cos\theta - ri\sin\theta = r(\cos\theta - i\sin\theta).$$

Because $\cos\theta = \cos(-\theta)$ and $\sin\theta = -\sin(-\theta)$, we have

$$\overline{z} = r(\cos\theta - i\sin\theta) = r(\cos(-\theta) - i(-\sin(-\theta))) = r(\cos(-\theta) + i\sin(-\theta)) = re^{-i\theta}.$$

□

Important: If $z = re^{i\theta}$, then $\overline{z} = re^{-i\theta}$.

Problem 7.14:

(a) Use exponential form to prove that $z\overline{z} = |z|^2$ for all complex numbers z.

(b) Use exponential form to prove that $|zw| = |z||w|$ for all complex numbers z and w.

(c) Prove that complex number multiplication is associative. In other words, show that if w_1, w_2, and w_3 are complex numbers, then $(w_1w_2)w_3 = w_1(w_2w_3)$.

Solution for Problem 7.14:

(a) Let $z = re^{i\theta}$, so $|z| = r$ and $\overline{z} = re^{-i\theta}$. Therefore, we have

$$z\overline{z} = (re^{i\theta})(re^{-i\theta}) = r^2e^{i\theta - i\theta} = r^2 = |z|^2.$$

(b) Back on page 205, we bashed through this proof with rectangular form, but exponential form gives the result immediately. Let $z = re^{i\theta}$ and $w = se^{i\phi}$. Then, $|z| = r$, $|w| = s$, and

$$|zw| = |(re^{i\theta})(se^{i\phi})| = |rse^{i(\theta+\phi)}| = |rs||e^{i(\theta+\phi)}| = rs = |z||w|.$$

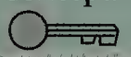
(c) Let $w_1 = r_1e^{i\theta_1}$, $w_2 = r_2e^{i\theta_2}$, and $w_3 = r_3e^{i\theta_3}$. Then, we have

$$(w_1w_2)w_3 = \left(r_1e^{i\theta_1} \cdot r_2e^{i\theta_2}\right) r_3e^{i\theta_3} = r_1r_2e^{i(\theta_1+\theta_2)} \cdot r_3e^{i\theta_3} = r_1r_2r_3e^{i(\theta_1+\theta_2+\theta_3)},$$
$$w_1(w_2w_3) = r_1e^{i\theta_1}\left(r_2e^{i\theta_2} \cdot r_3e^{i\theta_3}\right) = r_1e^{i\theta_1}\left(r_2r_3e^{i(\theta_2+\theta_3)}\right) = r_1r_2r_3e^{i(\theta_1+\theta_2+\theta_3)},$$

so $(w_1w_2)w_3 = w_1(w_2w_3)$.

□

Concept: Exponential form sometimes offers an easy way to prove facts involving complex number multiplication.



Problem 7.15: Express $e^{ix} \cdot e^{iy}$ in terms of $\cos x$, $\sin x$, $\cos y$, and $\sin y$. How does your result make remembering the angle sum identities for $\cos(x + y)$ and $\sin(x + y)$ easy?

Solution for Problem 7.15: We have

$$\begin{aligned} e^{ix} \cdot e^{iy} &= (\cos x + i \sin x)(\cos y + i \sin y) \\ &= \cos x \cos y - \sin x \sin y + i(\sin x \cos y + \cos x \sin y). \end{aligned}$$

However, we also have

$$e^{ix} \cdot e^{iy} = e^{i(x+y)} = \cos(x + y) + i \sin(x + y).$$

Our two expressions for $e^{ix} \cdot e^{iy}$ must be equal:

$$\cos(x + y) + i \sin(x + y) = \cos x \cos y - \sin x \sin y + i(\sin x \cos y + \cos x \sin y).$$

Two complex numbers are equal if and only if their real parts are equal and their imaginary parts are equal. Therefore, the equation above gives us

$$\begin{aligned} \cos(x + y) &= \cos x \cos y - \sin x \sin y, \\ \sin(x + y) &= \sin x \cos y + \cos x \sin y. \end{aligned}$$

These are the familiar angle sum identities, and now we have a way to quickly remember them, rather than having to memorize them. □

We don't call our work above a "proof" of the angle sum identities because our explanation above depends upon the fact that the function $f(x) = \cos x + i \sin x$ behaves like an exponential function. However, we proved this behavior using the angle sum identities! So, using this behavior to "prove" the angle sum identities would be circular reasoning. (By "circular reasoning," we mean using a statement we wish to prove as a step in proving the statement itself. This is clearly invalid!)

Problem 7.16:

(a) Find $\cos \theta$ in terms of $e^{i\theta}$ and $e^{-i\theta}$.

(b) Explain how you can use your relationship from part (a) to remember the product-to-sum identity for $\cos x \cos y$.

Solution for Problem 7.16:

(a) We have

$$\begin{aligned} e^{i\theta} &= \cos \theta + i \sin \theta, \\ e^{-i\theta} &= \cos(-\theta) + i \sin(-\theta) = \cos \theta - i \sin \theta, \end{aligned}$$

where we have used the identities $\cos(-\theta) = \cos \theta$ and $\sin(-\theta) = -\sin \theta$ in our expression for $e^{-i\theta}$. Adding our equations for $e^{i\theta}$ and $e^{-i\theta}$ eliminates $i \sin \theta$ and gives $e^{i\theta} + e^{-i\theta} = 2 \cos \theta$, so

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}.$$

As an Exercise, you'll derive a similar expression for $\sin \theta$.

Important: For any angle θ, we have

$$\begin{aligned} \cos \theta &= \frac{e^{i\theta} + e^{-i\theta}}{2}, \\ \sin \theta &= \frac{e^{i\theta} - e^{-i\theta}}{2i}. \end{aligned}$$

(b) Using the identity from the previous part, we have

$$\begin{aligned}\cos x \cos y &= \left(\frac{e^{ix}+e^{-ix}}{2}\right)\left(\frac{e^{iy}+e^{-iy}}{2}\right)\\ &= \frac{1}{4}\left(e^{i(x+y)}+e^{i(x-y)}+e^{i(y-x)}+e^{i(-x-y)}\right)\\ &= \frac{1}{2}\left(\frac{e^{i(x+y)}+e^{-i(x+y)}}{2}+\frac{e^{i(x-y)}+e^{-i(x-y)}}{2}\right).\end{aligned}$$

Again applying the identity from part (a), we see that the two fractions in parentheses above equal $\cos(x+y)$ and $\cos(x-y)$, and we have the product-to-sum identity for a product of cosines:

$$\cos x \cos y = \frac{1}{2}\left(\cos(x+y)+\cos(x-y)\right).$$

□

Concept: The exponential form of complex numbers is a helpful way to remember trigonometric identities.

Problem 7.17: Find the positive integer n such that

$$\arctan\frac{1}{3}+\arctan\frac{1}{4}+\arctan\frac{1}{5}+\arctan\frac{1}{n}=\frac{\pi}{4}.$$

(Source: AIME)

Solution for Problem 7.17: We could tackle this problem by taking the tangent of both sides and using the angle sum identity for tangent repeatedly. But complex numbers offer a much nicer solution. On the left, we have a sum of angles. Where else have we seen a sum of angles? When we multiply complex numbers in polar or exponential form, the argument of the product is the sum of the arguments of the factors. So, we can view each term on the left as the argument of a complex number, and the sum on the left is the argument of the product of these complex numbers.

But what complex number has argument $\arctan\frac{1}{3}$? Comparing the rectangular form to the polar form quickly gives the answer, since $x+yi = r(\cos\theta + i\sin\theta)$ gives us $x = r\cos\theta$ and $y = r\sin\theta$. Dividing these gives us $\frac{y}{x} = \frac{\sin\theta}{\cos\theta} = \tan\theta$. So, a complex number with $\tan\theta = \frac{1}{3}$ is $x+yi = 3+1i$. Similarly, $4+1i$ has argument $\arctan\frac{1}{4}$, $5+1i$ has argument $\arctan\frac{1}{5}$, and $n+1i$ has argument $\arctan\frac{1}{n}$. The argument of the product

$$(3+i)(4+i)(5+i)(n+i)$$

is equal to the sum of the arguments of these four complex numbers. From the expressions we just found for the arguments of these four numbers, we see that the argument of this product is

$$\arctan\frac{1}{3}+\arctan\frac{1}{4}+\arctan\frac{1}{5}+\arctan\frac{1}{n},$$

which we are given is equal to $\frac{\pi}{4}$. Since the argument of $(3+i)(4+i)(5+i)(n+i)$ equals $\frac{\pi}{4}$, and $\cos\frac{\pi}{4} = \sin\frac{\pi}{4}$, the real and imaginary parts of the product must be equal positive numbers. Expanding the product gives

$$\begin{aligned}(3+i)(4+i)(5+i)(n+i) &= (11+7i)(5+i)(n+i)\\ &= (48+46i)(n+i)\\ &= (48n-46)+(48+46n)i.\end{aligned}$$

We therefore have $48n - 46 = 48 + 46n$, from which we find $n = 47$. □

Exercises

7.2.1 Express each of the following in exponential form:

(a) $-32i$

(b) $-8 + 8i\sqrt{3}$

(c) $-\frac{\sqrt{2}}{6} + \frac{\sqrt{2}}{6}i$

(d) $2\sqrt{15} + 2i\sqrt{5}$

7.2.2 Express each of the following in rectangular form:

(a) $e^{4\pi i}$

(b) $e^{5\pi i/4}$

(c) $e^{-5\pi i/3}$

(d) $e^{17\pi i/3}$

7.2.3 Prove that $\sin\theta = \dfrac{e^{i\theta} - e^{-i\theta}}{2i}$.

7.2.4 Reproduce the angle difference identities for sine and cosine by expressing $\dfrac{e^{ix}}{e^{iy}}$ in rectangular form.

7.2.5 Compute the following in your head. (Don't use paper, pencil, pen, or a calculator.)

(a) $\dfrac{8 + 8i}{-3 + 3i}$

(b) $\dfrac{-12 - 4i\sqrt{3}}{3 + 3i\sqrt{3}}$

7.2.6 Evaluate each of the following:

(a) $(1 - i)^{16}$

(b) $(-2\sqrt{3} + 6i)^8$

7.2.7 For what positive integers n is $(-3 + i\sqrt{3})^n$ a real number?

7.2.8 Show that $e^{i\alpha} + e^{i\beta} = 2\cos\left(\frac{\alpha-\beta}{2}\right)e^{i(\alpha+\beta)/2}$.

7.3 Roots of Unity

In this section, we focus on the roots of a special class of polynomials, those of the form

$$f(x) = x^n - 1,$$

where n is a positive integer. The roots of such a polynomial are called the **n^{th} roots of unity**. We call these the roots of unity because finding the roots of $x^n - 1$ is equivalent to finding those numbers whose n^{th} power is 1. Obviously, 1 is one of these roots, but if $n > 1$, then there are other roots of f, as well. We explore these other roots in this section because they appear in so many areas of mathematics.

Problems

Problem 7.18: Find the roots of each of the following polynomials in rectangular and exponential forms. Plot the roots of each on the complex plane.

(a) $f(x) = x^3 - 1$

(b) $g(x) = x^4 - 1$

(c) $h(x) = x^6 - 1$

Problem 7.19: In this problem, we find the roots of $f(x) = x^5 - 1$.

(a) How many roots does f have? Why is it difficult to find the roots of f using the techniques you used in the previous problem?

(b) Let $t = re^{i\theta}$. If $f(t) = 0$, then what is r?

(c) If $f(t) = 0$, then what are the possible values of t?

Problem 7.20: Let n be a positive integer. What are the roots of $f(x) = x^n - 1$? If we plot the roots in the complex plane, then they are the vertices of what type of figure?

Problem 7.21: Suppose ω is a root of $f(z) = z^n - 1$. Which of the following *must* also be a root of $f(z)$?

(a) $\overline{\omega}$ (b) $\omega - 1$ (c) $\dfrac{1}{\omega}$

Problem 7.22: In this problem, we extend our understanding of the roots of unity to find the roots of other complex numbers.

(a) At most how many different complex numbers have fourth powers equal to $-8 - 8i\sqrt{3}$?

(b) Find all complex numbers whose fourth power is $-8 - 8i\sqrt{3}$. Plot all of them in the complex plane.

(c) How are your answers to part (b) related to the fourth roots of unity?

(d) Let $x = re^{i\theta}$. How are the n^{th} roots of x related to the n^{th} roots of unity?

Problem 7.23: An n^{th} root of unity, z, is called a **primitive n^{th} root of unity** if $z^n = 1$ and $z^k \neq 1$ for $k = 1, 2, 3, \ldots, n-1$. For example, -1 is a primitive square root of unity, since $(-1)^2 = 1$, but $(-1)^1 \neq 1$.

(a) Find the primitive cube roots of unity.

(b) Find the primitive fourth roots of unity.

(c) Find the primitive sixth roots of unity.

(d) Let n be a positive integer greater than 1. Prove that $e^{2\pi i/n}$ is a primitive n^{th} root of unity.

Problem 7.24: Is $e^{6\pi i/5}$ a primitive fifth root of unity?

Problem 7.25: Let $\omega = e^{2\pi ki/n}$, where k is a positive integer, be an n^{th} root of unity.

(a) Show that if $\gcd(k, n) = 1$, then ω is a primitive n^{th} root of unity.

(b) Show that if $\gcd(k, n) \neq 1$, then ω is not a primitive n^{th} root of unity.

(c) Let a and b be positive integers such that $0 < b < a < n$. Show that if ω is a primitive n^{th} root of unity, then it is impossible to have $\omega^a = \omega^b$. Why does this show that the first $n - 1$ powers of ω are the $n - 1$ n^{th} roots of unity besides 1?

Problem 7.18: Find the roots of each of the following polynomials in rectangular and exponential forms. Plot the roots of each on the complex plane.

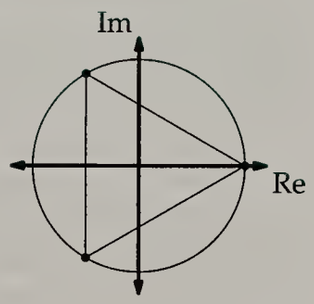
(a) $f(x) = x^3 - 1$

(b) $g(x) = x^4 - 1$

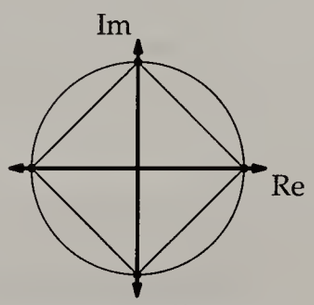
(c) $h(x) = x^6 - 1$

Solution for Problem 7.18:

(a) Factoring $x^3 - 1$ as a difference of cubes, we have $f(x) = (x-1)(x^2+x+1)$. (We also could have noticed that $f(1) = 0$, so $x - 1$ is a factor of f.) Therefore, 1 is a root of f, and so are the roots of the quadratic $x^2 + x + 1$. Applying the quadratic formula gives us the roots $x = \frac{-1 \pm i\sqrt{3}}{2}$. So, the roots in rectangular form are $1, -\frac{1}{2} + \frac{\sqrt{3}}{2}i, -\frac{1}{2} - \frac{\sqrt{3}}{2}i$, and in exponential form, they are $e^0, e^{2\pi i/3}, e^{4\pi i/3}$. Since the magnitude of each root is 1, we can plot each by plotting the terminal point of the angle of the exponential form of the root. Connecting the three points produces an equilateral triangle.



(b) Factoring $x^4 - 1$ as a difference of squares gives $(x^2-1)(x^2+1)$. The roots of $x^2 - 1$ are 1 and -1, and the roots of $x^2 + 1$ are i and $-i$, so the roots of $f(x)$ in rectangular form are $1, i, -1, -i$. In exponential form, these roots are $e^0, e^{\pi i/2}, e^{\pi i}, e^{3\pi i/2}$. Again, the magnitude of each root is 1, so we can plot each root by locating the terminal point of its angle in exponential form. As shown at right, plotting the roots on the complex plane produces the vertices of a square.



(c) Factoring $x^6 - 1$ as a difference of squares gives $(x^3-1)(x^3+1)$. Factoring each of these as a difference or sum of cubes gives us

$$f(x) = x^6 - 1 = (x^3-1)(x^3+1) = (x-1)(x^2+x+1)(x+1)(x^2-x+1).$$

So, 1 and -1 are roots of f, and applying the quadratic formula to the two quadratics tells us that the following are the roots of f in rectangular form:

$$1, \frac{-1+i\sqrt{3}}{2}, \frac{-1-i\sqrt{3}}{2}, -1, \frac{1+i\sqrt{3}}{2}, \frac{1-i\sqrt{3}}{2}.$$

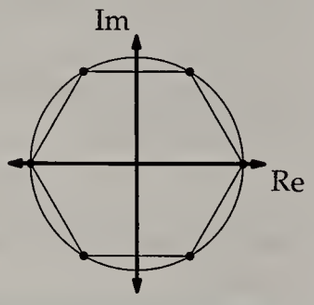
In exponential form, these are

$$e^0, e^{2\pi i/3}, e^{4\pi i/3}, e^{\pi i}, e^{\pi i/3}, e^{5\pi i/3}.$$

Putting these in order by increasing argument (angle), we have

$$e^0, e^{\pi i/3}, e^{2\pi i/3}, e^{\pi i}, e^{4\pi i/3}, e^{5\pi i/3}.$$

Each argument is $\pi/3$ more than the previous, and we aren't surprised to find that these roots are evenly spaced about the unit circle. Connecting them in this order gives us a regular hexagon, as shown at right.



□

Problem 7.19: In this problem, we find the roots of $f(x) = x^5 - 1$.

(a) How many roots does f have? Why is it difficult to find the roots of f using the techniques we used in the previous problem?

(b) Let $t = re^{i\theta}$. If $f(t) = 0$, then what is r?

(c) If $f(t) = 0$, then what are the possible values of t?

Solution for Problem 7.19:

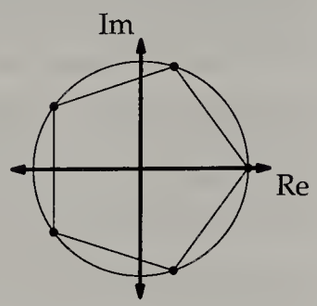
(a) Because f has degree 5, it has 5 roots. We can't use the difference of squares or difference of cubes factorizations, but we can see that $f(1) = 0$, so 1 is a root. Factoring out $x - 1$ gives

$$f(x) = (x-1)(x^4 + x^3 + x^2 + x + 1).$$

Unfortunately, now we're stuck. None of our typical root-finding techniques helps us with $x^4+x^3+x^2+x+1$. We'll have to try something different.

(b) Having seen that all the roots of x^3-1, x^4-1, and x^6-1 are of the form $e^{i\theta}$, we try to find roots in exponential form. We let $t = re^{i\theta}$, where r is a nonnegative real number and t is a root of f. Since $f(t) = 0$, we have $\left(re^{i\theta}\right)^5 = 1$, or $r^5e^{5i\theta} = 1$. The magnitude of the right side is 1, so the magnitude of the left side must be 1. Since r is nonnegative, we have $\left|r^5e^{5i\theta}\right| = r^5$, so now we have $r^5 = 1$, which means $r = 1$. (To see why there are no other nonnegative real numbers such that $r^5 = 1$, note that $r^5 > 1$ if $r > 1$ and $r^5 < 1$ if $0 \le r < 1$.)

(c) Since $f(t) = 0$ and $t = re^{i\theta}$ gave us $r^5e^{5i\theta} = 1$, and we found that $r = 1$, we must have $e^{5i\theta} = 1$. As we saw in Problem 7.12, if $e^{5i\theta} = 1$, then 5θ must be an integer multiple of 2π. Therefore, we must have $5\theta = 2\pi k$ for some integer k in order to have $e^{5i\theta} = 1$. Dividing by 5 gives us $\theta = 2\pi k/5$, so the roots are of the form $t = e^{2\pi ki/5}$. Letting k be each of 0, 1, 2, 3, and 4 gives us the roots of f:



$$e^0, e^{2\pi i/5}, e^{4\pi i/5}, e^{6\pi i/5}, e^{8\pi i/5}.$$

Any other integer k just reproduces one of these values. For example, if $k = 5$, then $\theta = 2\pi k/5 = 2\pi$, which produces the same root as $\theta = 0$. Similarly, increasing k by 5 always increases θ by 2π, which doesn't change $e^{i\theta}$ because $e^{i(\theta+2\pi)} = e^{i\theta} \cdot e^{2\pi i} = e^{i\theta} \cdot 1 = e^{i\theta}$.

We aren't at all surprised to see that the roots of f are equally spaced around the unit circle, and that connecting them forms a regular pentagon.

□

Problem 7.20: Let n be a positive integer. What are the roots of $f(x) = x^n - 1$? If we plot the roots in the complex plane, then they are the vertices of what type of figure?

Solution for Problem 7.20: The previous two problems give us a clear guide. We let t be a root of $f(x)$, and write t in exponential form: $t = re^{i\theta}$, where $r \ge 0$. Since $f(t) = 0$, we must have $(re^{i\theta})^n = 1$, so $r^ne^{in\theta} = 1$. This tells us that $\left|r^ne^{in\theta}\right| = 1$. Since $r \ge 0$, we have $\left|r^ne^{in\theta}\right| = r^n$, so $\left|r^ne^{in\theta}\right| = 1$ means that $r^n = 1$, which gives us $r = 1$. (Yes, this should all look familiar—it's exactly the same reasoning we used to find the roots of $x^5 - 1$.)

Since $r = 1$, the equation $f(t) = 1$ is now $e^{in\theta} = 1$, which means we must have $n\theta = 2\pi k$ for some integer k. Dividing by n gives us $\theta = 2\pi k/n$, so our roots are of the form $t = e^{2k\pi i/n}$. Letting k be each of $0, 1, 2, \ldots, n-1$ gives us the roots of f:

$$e^0, e^{2\pi i/n}, e^{4\pi i/n}, e^{6\pi i/n}, \ldots, e^{2(n-1)\pi i/n},$$

Notice that any other integer k just reproduces one of these values. For example, if $k = n$, then $\theta = 2\pi k/n = 2\pi$, which produces the same root as $\theta = 0$. Similarly, increasing k by n always increases θ by 2π, which doesn't change $e^{i\theta}$.

These n roots are equally spaced about the unit circle. The angle of each is $2\pi/n$ from its neighbors, including the roots e^0 and $e^{2(n-1)\pi i/n}$. Therefore, plotting these n roots on the complex plane produces the vertices of a regular polygon with n sides. □

Important: The n roots of

$$f(x) = x^n - 1$$

are

$$e^0, e^{2\pi i/n}, e^{4\pi i/n}, e^{6\pi i/n}, \ldots, e^{2(n-1)\pi i/n}.$$

These are also referred to as the n^{th} **roots of unity.**

Problem 7.21: Suppose ω is a root of $f(z) = z^n - 1$. Which of the following *must* also be a root of $f(z)$?

(a) $\overline{\omega}$ (b) $\omega - 1$ (c) $\dfrac{1}{\omega}$

Solution for Problem 7.21:

(a) As we showed on page 212, if z is a root of a polynomial f with real coefficients, then $\overline{z}$ is also a root f. Therefore if ω is a root of $f(z) = z^n - 1$, then so is $\overline{\omega}$. In other words, if ω is an n^{th} root of unity, then so is $\overline{\omega}$.

(b) We can quickly see that $\omega - 1$ is not necessarily a root of f by noting that $\omega = 1$ is a root of $f(z) = z^n - 1$, but this gives $\omega - 1 = 0$, which is clearly not a root of f.

(c) Since $\omega^n = 1$, we have

$$f\left(\frac{1}{\omega}\right) = \left(\frac{1}{\omega}\right)^n - 1 = \frac{1}{\omega^n} - 1 = \frac{1}{1} - 1 = 0.$$

Therefore, $\frac{1}{\omega}$ is a root of f.

□

Now that we know how to find n^{th} roots of unity, we can find the n^{th} roots of other complex numbers.

Problem 7.22:

(a) At most how many different complex numbers have fourth powers equal to $-8 - 8i\sqrt{3}$?

(b) Find all complex numbers whose fourth power is $-8 - 8i\sqrt{3}$. Plot all of them in the complex plane.

(c) How are your answers to part (b) related to the fourth roots of unity?

(d) Let $x = re^{i\theta}$. How are the n^{th} roots of x related to the n^{th} roots of unity?

Solution for Problem 7.22:

(a) We seek the complex numbers z such that $z^4 = -8 - 8i\sqrt{3}$. Rewriting this as $z^4 + 8 + 8i\sqrt{3} = 0$, we see that we seek the roots of the polynomial $z^4 + 8 + 8i\sqrt{3}$. Since this polynomial has degree 4, it has 4 roots. So, there are no more than 4 different complex numbers whose fourth power is $-8 - 8i\sqrt{3}$. (We don't know for sure yet that there are 4, because we haven't shown that the polynomial doesn't have any repeated roots.)

(b) We found the roots of unity by using the exponential form of complex numbers. Let's try the same here. We let $z = re^{i\theta}$, where $r \geq 0$, be a solution to the equation $z^4 = -8 - 8i\sqrt{3}$, so we have

$$r^4 e^{4i\theta} = -8 - 8i\sqrt{3}.$$

It's not so clear what to do with that, so we write the right side in exponential form, too. This gives us

$$r^4 e^{4i\theta} = 16\left(-\frac{1}{2} - \frac{\sqrt{3}}{2}i\right) = 16e^{4\pi i/3}.$$

Comparing the magnitudes of both sides of $r^4 e^{4i\theta} = 16e^{4\pi i/3}$ gives us $r^4 = 16$, which gives us $r = 2$ (remember, r is a nonnegative real number). All that's left is to find θ such that $e^{4i\theta} = e^{4\pi i/3}$. Clearly, $\theta = \frac{\pi}{3}$ works, and gives us the solution $z = re^{\pi i/3} = 2\left(\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) = 1 + i\sqrt{3}$; however, we expect to find four solutions, not just one. What are the others?

We find the others by recalling that $e^{i(\alpha+2\pi)} = e^{i\alpha}$ for all angles α. So,

$$e^{4\pi i/3} = e^{i[(4\pi/3)+2\pi]} = e^{i[(4\pi/3)+4\pi]} = e^{i[(4\pi/3)+6\pi]} = \cdots.$$

This means that we can write the equation $e^{4i\theta} = e^{4\pi i/3}$ as $e^{4i\theta} = e^{i[(4\pi/3)+2\pi k]}$, where k is an integer. From this equation, we have $4\theta = \frac{4\pi}{3} + 2\pi k$. Dividing by 4 gives $\theta = \frac{\pi}{3} + \frac{\pi k}{2}$. When $k = 0$, we have the first solution we found, $\theta = \frac{\pi}{3}$. Letting $k = 1, 2, 3$ gives $\theta = \frac{5\pi}{6}, \frac{4\pi}{3}, \frac{11\pi}{6}$, respectively. When $k = 4$, we have $\theta = \frac{\pi}{3} + 2\pi$, which gives us the same value of $re^{i\theta}$ that $k = 0$ gave us. Similarly, any other value of k reproduces one of the four solutions we have already found, since increasing k by 4 increases θ by 2π, which leaves $e^{i\theta}$ unchanged.

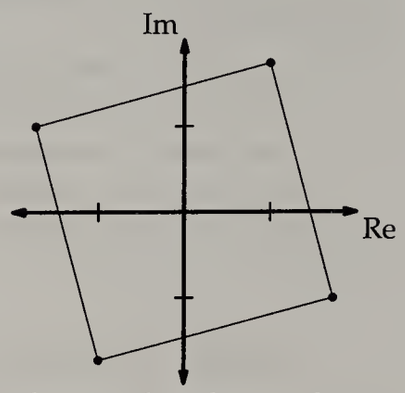
Putting our four values of θ together with $r = 2$ gives us the following four numbers whose fourth power is $-8 - 8i\sqrt{3}$:

$$2e^{\pi i/3}, 2e^{5\pi i/6}, 2e^{4\pi i/3}, 2e^{11\pi i/6}.$$

In rectangular form, these are

$$1 + i\sqrt{3},\ -\sqrt{3} + i,\ -1 - i\sqrt{3},\ \sqrt{3} - i.$$

We have plotted these in the complex plane at right. Connecting them forms a square.



(c) Seeing that the four fourth roots of $-8 - 8i\sqrt{3}$ are the vertices of a square in the complex plane makes us wonder if these roots are related to the roots of unity. In our solution to part (b), we found the desired values of θ with the equation $\theta = \frac{\pi}{3} + \frac{\pi k}{2}$. The angles $\frac{\pi k}{2}$ for $k = 0, 1, 2, 3$ are the angles of the fourth roots of unity. This allows us to express each of our solutions for $re^{i\theta}$ in terms of one of the fourth roots of unity:

$$\begin{aligned} 2e^{(\pi i/3)+0} &= 2e^{\pi i/3} \cdot e^{0i}, \\ 2e^{(\pi i/3)+(\pi i/2)} &= 2e^{\pi i/3} \cdot e^{\pi i/2}, \\ 2e^{(\pi i/3)+(\pi i)} &= 2e^{\pi i/3} \cdot e^{\pi i}, \\ 2e^{(\pi i/3)+(3\pi i/2)} &= 2e^{\pi i/3} \cdot e^{3\pi i/2}. \end{aligned}$$

Aha! To produce the solutions of $z^4 = 16e^{4\pi i/3}$, we multiply each of the fourth roots of unity by $2e^{\pi i/3}$.

(d) Part (c) gives us a guide for generating the n^{th} roots of any complex number $re^{i\theta}$. First, we find the n^{th} roots of unity, which are $e^{2\pi ik/n}$ for $0 \le k < n$. Next, we note that if $z^n = re^{i\theta}$, then raising both sides to the power $\frac{1}{n}$ gives $z = \sqrt[n]{r}e^{i\theta/n}$ as one solution. We then multiply each of the n^{th} roots of unity by this number to produce the n roots of $re^{i\theta}$.

To confirm that this procedure does indeed produce the n values of z such that $z^n = re^{i\theta}$, let $z = \sqrt[n]{r}e^{i\theta/n} \cdot \omega$, where ω is one of the n^{th} roots of unity. Since ω is an n^{th} root of unity, we have $\omega^n = 1$, and

$$z^n = \left(\sqrt[n]{r}e^{i\theta/n} \cdot \omega\right)^n = \left(\sqrt[n]{r}\right)^n \cdot \left(e^{i\theta/n}\right)^n \cdot \omega^n = re^{i\theta},$$

so z is indeed one of the n^{th} roots of $re^{i\theta}$. Since no two of the n^{th} roots of unity have the same argument, multiplying each of these by $\sqrt[n]{r}e^{i\theta/n}$ produces n complex numbers that have n different arguments. In other words, no two of these n numbers are the same. Therefore, we have found the n different numbers whose n^{th} power is $re^{i\theta}$.

□

Problem 7.23: An n^{th} root of unity, z, is called a **primitive n^{th} root of unity** if $z^n = 1$ and $z^k \neq 1$ for $k = 1, 2, 3, \ldots, n-1$. For example, -1 is a primitive square root of unity, since $(-1)^2 = 1$, but $(-1)^1 \neq 1$.

(a) Find the primitive cube roots of unity.

(b) Find the primitive fourth roots of unity.

(c) Find the primitive sixth roots of unity.

(d) Let n be a positive integer greater than 1. Prove that $e^{2\pi i/n}$ is a primitive n^{th} root of unity.

Solution for Problem 7.23:

(a) By the definition of primitive roots of unity, if ω is a primitive cube root of unity, then $\omega^3 = 1$, $\omega \neq 1$, and $\omega^2 \neq 1$. The three cube roots of unity are 1, $e^{2\pi i/3}$, and $e^{4\pi i/3}$. Clearly the first is not a primitive cube root of unity since we must have $\omega \neq 1$. We check the other two by computing their squares. We have

$$\left(e^{2\pi i/3}\right)^2 = e^{4\pi i/3} \neq 1,$$
$$\left(e^{4\pi i/3}\right)^2 = e^{8\pi i/3} \neq 1.$$

Therefore, both $e^{2\pi i/3}$ and $e^{4\pi i/3}$ are primitive cube roots of unity.

(b) The fourth roots of unity are 1, $e^{\pi i/2}$, $e^{\pi i}$, $e^{3\pi i/2}$. The first and third are not primitive fourth roots because the first equals 1 and the square of $e^{\pi i}$ is 1. The other two are primitive fourth roots because for each one, the root, its square, and its cube all do not equal 1. So, the primitive fourth roots of unity are $e^{\pi i/2}$ and $e^{3\pi i/2}$.

(c) The sixth roots of unity are

$$1, e^{\pi i/3}, e^{2\pi i/3}, e^{\pi i}, e^{4\pi i/3}, e^{5\pi i/3}.$$

The first is not a primitive sixth root. Both $e^{2\pi i/3}$ and $e^{4\pi i/3}$ are cube roots of unity, so they are not primitive sixth roots. Similarly, $e^{\pi i}$ is a square root of unity, so it's not a primitive sixth root. If $\omega = e^{\pi i/3}$, then $\omega^k \neq 1$ for $1 \leq k \leq 5$ (since ω^k for these values represent the 5 sixth roots of unity besides 1). To check the powers of $e^{5\pi i/3}$, we let $\omega = e^{5\pi i/3}$ and compute the first 5 powers of ω. When the argument of the result is greater than 2π, we can subtract integer multiples of 2π without changing the result. We find:

$$\omega^1 = e^{5\pi i/3},$$
$$\omega^2 = e^{10\pi i/3} = e^{4\pi i/3},$$
$$\omega^3 = e^{15\pi i/3} = e^{3\pi i/3},$$
$$\omega^4 = e^{20\pi i/3} = e^{2\pi i/3},$$
$$\omega^5 = e^{25\pi i/3} = e^{1\pi i/3}.$$

Not only are none of these five powers of ω equal to 1, but all five are different. Interesting! We now know that $e^{5\pi i/3}$ is a primitive cube root of unity, and we have a clue that we'll investigate in Problem 7.25 for determining which n^{th} roots of unity are primitive.

(d) In each of the previous parts, we saw that $e^{2\pi i/n}$ is an n^{th} root of unity. We can quickly prove that this is always the case by noting that if $\omega = e^{2\pi i/n}$, then $\omega^k = e^{2\pi ki/n}$. If $0 < k < n$, then the argument of ω^k is between, but not equal to, 0 and 2π. Therefore, ω^k cannot equal 1. So, we conclude that $\omega = e^{2\pi i/n}$ is a primitive n^{th} root of unity.

□

The first three parts above suggest a pattern regarding which roots are primitive and which are not. When $n = 3$, we found that $e^{2\pi ki/n}$ is a primitive n^{th} root of unity for $k = 1$ and $k = 2$. When $n = 4$, we found that $e^{2\pi ki/n}$ is a primitive n^{th} root of unity for $k = 1$ and $k = 3$. When $n = 6$, we found that $e^{2\pi ki/n}$ is a primitive n^{th} root of unity for $k = 1$ and $k = 5$. We might guess that $e^{2\pi ki/n}$ is a primitive n^{th} root of unity if $k = 1$ or $k = n - 1$, and not a primitive root otherwise. Before we dive into proving that, let's take a look at a fifth root of unity.

Problem 7.24: Is $e^{6\pi i/5}$ a primitive fifth root of unity?

Solution for Problem 7.24: We let $\omega = e^{6\pi i/5}$ and find the first four powers of ω. As before, we can subtract positive

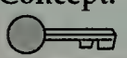
integer multiples of 2π from the argument of each result without changing its value:

$$\begin{aligned}\omega^1 &= e^{6\pi i/5},\\ \omega^2 &= e^{12\pi i/5} = e^{2\pi i/5},\\ \omega^3 &= e^{18\pi i/5} = e^{8\pi i/5},\\ \omega^4 &= e^{24\pi i/5} = e^{4\pi i/5}.\end{aligned}$$

Since none of these first four powers of ω equals 1, but $\omega^5 = 1$, we know that $\omega = e^{6\pi i/5}$ is a primitive fifth root of unity. Also, we see that these four powers of ω produce the four fifth roots of unity besides 1. □

The result of Problem 7.24 disproves our hypothesis that only $e^{2\pi i/n}$ and $e^{2\pi(n-1)i/n}$ can be primitive n^{th} roots of unity, since we found that $e^{2\pi ki/5}$ is a primitive fifth root of unity when $k = 3$. Back to the drawing board. Let's look back at our results above and see if we can find some other pattern in which of the n^{th} roots of unity are primitive n^{th} roots of unity.

In looking over the results of the past two problems, we see that $e^{2\pi ki/n}$ is a primitive n^{th} root of unity if k and n have no common divisors besides 1, and the root is not primitive otherwise. If we test all the eighth and ninth roots of unity, we'll find that this pattern holds. Moreover, we'll see the same pattern we saw when we discovered that the first four powers of $e^{6\pi i/5}$ produce the four 5^{th} roots of unity besides 1, and that the first five powers of $e^{5\pi i/3}$ produce the five 6^{th} roots of unity besides 1. These observations give us a strong clue how to quickly determine which n^{th} roots of unity are primitive.



Concept: Investigating examples can reveal patterns that help you conjecture and prove general results.

We use the notation $\gcd(p, q)$ to refer to the greatest common divisor of p and q.

Problem 7.25: Let $\omega = e^{2\pi ki/n}$, where k is a positive integer, be an n^{th} root of unity.

(a) Show that if $\gcd(k, n) = 1$, then ω is a primitive n^{th} root of unity.

(b) Show that if $\gcd(k, n) \neq 1$, then ω is not a primitive n^{th} root of unity.

(c) Let a and b be positive integers such that $0 < b < a < n$. Show that if ω is a primitive n^{th} root of unity, then it is impossible to have $\omega^a = \omega^b$. Why does this show that the first $n - 1$ powers of ω are the $n - 1$ n^{th} roots of unity besides 1?

Solution for Problem 7.25:

(a) Suppose that q is a positive integer such that $\omega^p = 1$, which means $e^{2\pi kpi/n} = 1$. Therefore, $2\pi kpi/n$ must be an integer multiple of $2\pi i$, so kp/n is an integer. Since k and n have no common factors besides 1, the only way kp/n can be an integer is if n divides p evenly. Therefore, we can only have $\omega^p = 1$ if p is a multiple of n. This means that there is no p with $1 \leq p \leq n - 1$ such that $\omega^p = 1$, so ω is a primitive n^{th} root of unity.

(b) We start just as in the previous part, with p being a positive integer such that $\omega^p = 1$, which means $e^{2\pi kpi/n} = 1$. Therefore, $2\pi kpi/n$ must be an integer multiple of $2\pi i$, so kp/n is an integer. However, if $\gcd(k, n) \neq 1$, then some factors of n cancel with factors of k in the fraction kp/n, which means there is an equivalent fraction with a smaller denominator. When p equals this denominator, the fraction is an integer. Therefore, we can have a value of p smaller than n such that $\omega^p = 1$.

To make this argument more formal, we can let $d = \gcd(k, n)$. Since d divides n and k, we have $n = dt$ and $k = du$ for some positive integers t and u, so $k/n = u/t$. We can now write ω as $\omega = e^{2\pi ki/n} = e^{2\pi ui/t}$. Now, when we raise ω to the t power, we have:

$$\omega^t = \left(e^{2\pi ui/t}\right)^t = e^{2\pi ui}.$$

Since u is an integer, we have $e^{2\pi ui} = 1$, which means $\omega^t = 1$. Moreover, since $n = dt$ and $d > 1$, we have $0 < t < n$, so $\omega^t = 1$ tells us that ω is not a primitive n^{th} root of unity.

(c) If $\omega^a = \omega^b$, then dividing both sides by ω^b gives $\omega^{a-b} = 1$. Since $\omega^{a-b} = \left(e^{2\pi ki/n}\right)^{a-b} = e^{2\pi k(a-b)i/n}$, we must have

$$e^{2\pi k(a-b)i/n} = 1.$$

Therefore, $2\pi k(a-b)i/n$ must be an integer multiple of $2\pi i$, which means $k(a-b)/n$ must be an integer. However, since ω is a primitive n^{th} root, we must have $\gcd(k,n) = 1$. So, if $k(a-b)/n$ is an integer, then n must divide $a-b$. But we have $0 < b < a < n$, which means that $0 < a-b < n$, so it is impossible for n to divide $a-b$. We have reached a contradiction, which means that our original assumption, $\omega^a = \omega^b$, must be false. We conclude that no two of the first $n-1$ powers of ω are equal.

Moreover, we can conclude that the first $n-1$ powers of any primitive n^{th} root of unity give us the $n-1$ n^{th} roots of unity besides 1. To see why, we note that because ω is an n^{th} root of unity, we have $\omega^n = 1$, so

$$(\omega^p)^n = \omega^{pn} = (\omega^n)^p = 1^p = 1,$$

which means ω^p is an n^{th} root of unity for any p. Because ω is primitive, none of the first $n-1$ powers of ω equals 1 and no two of the first $n-1$ powers of ω are equal. Therefore, these first $n-1$ powers of ω are the n^{th} roots of unity besides 1.

□

Important: For any positive integer n greater than 1, we have the following:

- The complex numbers

$$e^0, e^{2\pi i/n}, e^{4\pi i/n}, e^{6\pi i/n}, \ldots, e^{2(n-1)\pi i/n}$$

are the n different n^{th} roots of unity.

- If $\gcd(k,n) = 1$, then $\omega = e^{2\pi ik/n}$ is a primitive n^{th} root of unity, and

$$1, \omega, \omega^2, \omega^3, \ldots, \omega^{n-1}$$

are all of the n^{th} roots of unity.

Exercises

7.3.1 Find all x that satisfy $x^8 = 1$. Which solutions are primitive eighth roots of unity?

7.3.2 If the six solutions of $x^6 = -64$ are written in the form $a + bi$, where a and b are real, then find the product of the solutions with $a > 0$. *(Source: AHSME)*

7.3.3 Let ω be a sixth root of unity. Find all possible values of $\omega + \omega^2$.

7.3.4

(a) Show that each 12^{th} root of unity is a primitive k^{th} root of unity for some positive integer k. What are the possible values of k?

(b) For each integer k in part (a), count how many 12^{th} roots of unity are primitive k^{th} roots of unity.

7.3.5 Show that if ω is a primitive n^{th} root of unity for $n = 2m$, then ω^2 is a primitive m^{th} root of unity.

7.3.6

(a) Let α be a cube root of unity, and let β be a fifth root of unity. Prove that $\alpha\beta$ is a 15^{th} root of unity.

(b) Let α be a primitive cube root of unity, and let β be a primitive fifth root of unity. Prove that $\alpha\beta$ is a primitive 15^{th} root of unity.

(c) Let α be a primitive fourth root of unity, and let β be a primitive sixth root of unity. Is $\alpha\beta$ necessarily a primitive 24^{th} root of unity?

7.3.7★ We say that a polynomial is **irreducible over the rational numbers** if it cannot be factored as the product of polynomials with smaller positive degree and rational coefficients. Find a degree 4 polynomial with leading coefficient 1 and rational coefficients such that the polynomial is irreducible over the rational numbers and all the polynomial's roots are all twelfth roots of unity. *(Source: HMMT)* **Hints:** 182

7.4 Problems Involving Roots of Unity

Problems

Problem 7.26: In this problem, we find the roots of the polynomial $f(x) = x^8 + x^7 + x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$.

(a) What kind of series is the expression $x^8 + x^7 + x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$?

(b) Viewing $f(x)$ as a series of the type you determined in part (a), find another expression for $f(x)$.

(c) Find the roots of $f(x)$ by determining the values of x for which your expression in part (b) equals 0. (You can leave your answers in exponential form.)

Problem 7.27: Find the roots of $g(z) = z^6 + z^4 + z^2 + 1$.

Problem 7.28: Let $\omega = e^{4\pi i/7}$. Evaluate $(2+\omega)(2+\omega^2)(2+\omega^3)(2+\omega^4)(2+\omega^5)(2+\omega^6)$. **Hints:** 236, 252

Problem 7.29: Find all solutions to $z^6 + z^4 + z^3 + z^2 + 1 = 0$. *(Source: AIME)* **Hints:** 263

Problem 7.30: Find the sum of the n^{th} roots of unity.

Problem 7.31: Let n be a positive integer. In this problem, we evaluate the sum

$$\cos\frac{2\pi}{2n+1} + \cos\frac{2\cdot 2\pi}{2n+1} + \cos\frac{3\cdot 2\pi}{2n+1} + \cdots + \cos\frac{n\cdot 2\pi}{2n+1}.$$

(a) Evaluate the sum $\cos\frac{2\pi}{m} + \cos\frac{2\cdot 2\pi}{m} + \cos\frac{3\cdot 2\pi}{m} + \cdots + \cos\frac{m\cdot 2\pi}{m}$, where m is a positive integer.

(b) Evaluate the sum $\cos\frac{2\pi}{2n+1} + \cos\frac{2\cdot 2\pi}{2n+1} + \cos\frac{3\cdot 2\pi}{2n+1} + \cdots + \cos\frac{n\cdot 2\pi}{2n+1}$, where n is a positive integer.

Problem 7.32: The equation

$$x^{10} + (13x-1)^{10} = 0$$

has the 10 complex roots $r_1, \overline{r_1}, r_2, \overline{r_2}, r_3, \overline{r_3}, r_4, \overline{r_4}, r_5, \overline{r_5}$. Find the value of

$$\frac{1}{r_1\overline{r_1}} + \frac{1}{r_2\overline{r_2}} + \frac{1}{r_3\overline{r_3}} + \frac{1}{r_4\overline{r_4}} + \frac{1}{r_5\overline{r_5}}.$$

(Source: AIME)

Problem 7.26: Find the roots of the polynomial $f(x) = x^8 + x^7 + x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$.

Solution for Problem 7.26: If we're very experienced with polynomials, we might recognize $f(x)$ as a factor of $x^9 - 1$. But if we don't see this right way, all is not lost. The polynomial $f(x)$ is a geometric series; each term in the sum is produced by multiplying the previous term by a constant ratio. By writing $f(x)$ as

$$f(x) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6 + x^7 + x^8,$$

we see that we can think of $f(x)$ as a geometric series with first term 1 and constant ratio x. To simplify this sum, we multiply both sides by x, which gives

$$xf(x) = x + x^2 + x^3 + x^4 + x^5 + x^6 + x^7 + x^8 + x^9.$$

Subtracting our equation for $f(x)$ from our equation for $xf(x)$ gives $(x-1)f(x) = x^9 - 1$. What's wrong with this finish:

Bogus Solution:

Dividing both sides by $x - 1$ gives

$$f(x) = \frac{x^9 - 1}{x - 1}.$$

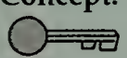
If $x^9 - 1 = 0$, then $f(x) = 0$. The solutions to $x^9 - 1 = 0$ are the ninth roots of unity, so the roots of f are the 9 ninth roots of unity.

Our first clue that this is incorrect is that f has degree 8, so it must have 8 roots, not 9. Where we went wrong is that the step from $(x-1)f(x) = x^9 - 1$ to $f(x) = \frac{x^9-1}{x-1}$ is only valid if $x \neq 1$. We must have $x \neq 1$ to avoid dividing by 0. When $x = 1$, our original expression for $f(x)$ gives us $f(1) = 9$. When $x \neq 1$, we do indeed have $f(x) = \frac{x^9-1}{x-1}$, so $f(x) = 0$ when $x \neq 1$ and $x^9 - 1 = 0$. Therefore, the solutions to $f(x) = 0$ are the ninth roots of unity *except for* $x = 1$. So, the roots of f are

$$e^{2\pi i/9}, e^{4\pi i/9}, e^{6\pi i/9}, e^{8\pi i/9}, e^{10\pi i/9}, e^{12\pi i/9}, e^{14\pi i/9}, e^{16\pi i/9}.$$

□

The key step in the solution to Problem 7.26 was noticing that the polynomial is a geometric series. This observation is particularly useful with polynomials of the form $x^n + x^{n-1} + x^{n-2} + \cdots + 1$.

Concept: 

Many problems involving polynomials of the form

$$x^n + x^{n-1} + x^{n-2} + \cdots + 1$$

can be solved by writing the polynomial in the form

$$\frac{x^{n+1} - 1}{x - 1},$$

and using our knowledge of roots of unity.

Problem 7.27: Find the roots of $g(z) = z^6 + z^4 + z^2 + 1$.

Solution for Problem 7.27: As in Problem 7.26, the polynomial is a geometric series. It doesn't have exactly the same form as the polynomial in Problem 7.26, but hopefully the same strategy will work. Multiplying $g(z) = z^6 + z^4 + z^2 + 1$ by z^2 gives

$$z^2 g(z) = z^8 + z^6 + z^4 + z^2.$$

Subtracting the equation for $g(z)$ from this gives $(z^2-1)g(z) = z^8 - 1$. If $z^2 \neq 1$, then we can divide both sides by $z^2 - 1$ to get

$$g(z) = \frac{z^8 - 1}{z^2 - 1}.$$

(We can also use the formula for summing geometric series to reach this expression for $g(z)$.) Therefore, we have $g(z) = 0$ if $z^8 = 1$ and $z^2 \neq 1$, which means that the roots of g are the eighth roots of unity except for 1 and -1. These are

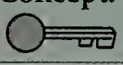
$$e^{\pi i/4}, e^{\pi i/2}, e^{3\pi i/4}, e^{5\pi i/4}, e^{3\pi i/2}, e^{7\pi i/4}.$$

□

Problem 7.28: Let $\omega = e^{4\pi i/7}$. Evaluate $(2+\omega)(2+\omega^2)(2+\omega^3)(2+\omega^4)(2+\omega^5)(2+\omega^6)$.

Solution for Problem 7.28: First, we recognize that ω is a seventh root of unity, because $\omega^7 = \left(e^{4\pi i/7}\right)^7 = e^{4\pi i} = 1$. Moreover, ω is a primitive seventh root of unity, so ω, ω^2, ω^3, ω^4, ω^5, and ω^6 are the six seventh roots of unity besides 1. But how does that help?

We could multiply the whole desired expression out. That will work, but it will take a fair amount of diligence and patience. Perhaps we can do something clever instead. Where else have we seen a product of binomials involving the seventh roots of unity?

Concept:  When working with a difficult expression, think about where else you have seen similar expressions.

We encounter the product of binomials involving the seventh roots of unity in the factored form of $z^7 - 1$:

$$z^7 - 1 = (z-1)(z-r_1)(z-r_2)(z-r_3)(z-r_4)(z-r_5)(z-r_6),$$

where the r_i are the seventh roots of unity besides 1. We've already seen that these roots are ω, ω^2, ω^3, ω^4, ω^5, and ω^6 (in some order), so we have

$$z^7 - 1 = (z-1)(z-\omega)(z-\omega^2)(z-\omega^3)(z-\omega^4)(z-\omega^5)(z-\omega^6).$$

Aha! We can make the desired expression appear by letting $z = -2$:

$$(-2)^7 - 1 = (-2-1)(-2-\omega)(-2-\omega^2)(-2-\omega^3)(-2-\omega^4)(-2-\omega^5)(-2-\omega^6).$$

Factoring -1 out of each of the final six expressions in parentheses gives us

$$-129 = -3(-1)^6(2+\omega)(2+\omega^2)(2+\omega^3)(2+\omega^4)(2+\omega^5)(2+\omega^6),$$

so

$$(2+\omega)(2+\omega^2)(2+\omega^3)(2+\omega^4)(2+\omega^5)(2+\omega^6) = \frac{129}{3} = 43.$$

□

Problem 7.29: Find all solutions to $z^6 + z^4 + z^3 + z^2 + 1 = 0$. *(Source: AIME)*

Solution for Problem 7.29: If only the z^6 were just z, then we could find the roots easily; the roots of $z^4 + z^3 + z^2 + z + 1$ are the fifth roots of unity besides 1:

$$z^4 + z^3 + z^2 + z + 1 = \frac{z^5 - 1}{z - 1}.$$

Aha! Adding $z^6 - z$ to both sides makes the left side the polynomial whose roots we seek, and we can factor the right side further:

$$\begin{aligned} z^6 + z^4 + z^3 + z^2 + 1 = \frac{z^5 - 1}{z - 1} + z^6 - z &= \frac{z^5 - 1}{z - 1} + z(z^5 - 1) \\ &= \frac{z^5 - 1 + z(z-1)(z^5 - 1)}{z - 1} \\ &= \frac{(z^2 - z + 1)(z^5 - 1)}{z - 1}. \end{aligned}$$

We can factor $z^5 - 1$ as $(z - 1)(z^4 + z^3 + z^2 + z + 1)$. The $z - 1$ cancels with the denominator and leaves

$$z^6 + z^4 + z^3 + z^2 + 1 = (z^2 - z + 1)(z^4 + z^3 + z^2 + z + 1).$$

The roots of $z^4 + z^3 + z^2 + z + 1$ are the fifth roots of unity besides 1. We could use the quadratic formula to find the roots of $z^2 - z + 1$, or we could note that $z^3 + 1 = (z + 1)(z^2 - z + 1)$, so the roots of $z^2 - z + 1$ are the nonreal cube roots of -1. Since $-1 = e^{\pi i}$, the nonreal cube roots of -1 are $e^{\pi i/3}$ and $e^{5\pi i/3}$.

Therefore, the solutions to the given equation are $e^{2\pi i/5}, e^{4\pi i/5}, e^{6\pi i/5}, e^{8\pi i/5}, e^{\pi i/3}, e^{5\pi i/3}$. □

Sidenote: In hindsight, it isn't surprising that the nonreal fifth roots of unity are also roots of $z^6 + z^4 + z^3 + z^2 + 1$. For any fifth root of unity, we have $z^5 = 1$, so $z^6 = z^5 \cdot z = z$. Therefore, if z is a nonreal fifth root of unity, we have

$$z^6 + z^4 + z^3 + z^2 + z + 1 = z + z^4 + z^3 + z^2 + 1,$$

which equals 0 because the nonreal fifth roots of unity are the roots of the polynomial $z^4 + z^3 + z^2 + z + 1$.

Problem 7.30: Find the sum of the n^{th} roots of unity.

Solution for Problem 7.30: First, we write out the sum:

$$e^{0\pi i/n} + e^{2\pi i/n} + e^{2\cdot 2\pi i/n} + e^{3\cdot 2\pi i/n} + e^{4\cdot 2\pi i/n} + \cdots + e^{(n-1)\cdot 2\pi i/n}.$$

Here are two ways to tackle this sum:

Method 1: It's a geometric series! This sum is an n-term geometric series with first term $e^{0\pi i/n}$ and common ratio $e^{2\pi i/n}$, so its sum is

$$\frac{e^{0\pi i/n} - e^{0\pi i/n} \cdot \left(e^{2\pi i/n}\right)^n}{1 - e^{2\pi i/n}} = \frac{1 - e^{2n\pi i/n}}{1 - e^{2\pi i/n}} = \frac{1 - e^{2\pi i}}{1 - e^{2\pi i/n}} = \frac{1 - 1}{1 - e^{2\pi i/n}} = 0.$$

Method 2: Consider the polynomial $z^n - 1$. The terms in the sum are the roots of $z^n - 1$, because they are the n^{th} roots of unity. We can see that the sum of the roots of this polynomial is 0 by writing the polynomial in the form

$$(z - r_1)(z - r_2)(z - r_3)\cdots(z - r_n),$$

where the r_i are the roots of the polynomial. To form each term in the expansion of this product, we choose one of the terms in each binomial, and then take the product of all the chosen terms. The full expansion includes every possible product we can form in this manner. For example, when we take $-r_1$ from $z - r_1$ and take z from each of the other $n - 1$ binomials, we get $-r_1 z^{n-1}$. Continuing in this way, the only way we can form a term with z^{n-1} is when we take $-r_i$ from one factor and z from each of the other $n - 1$ factors. Grouping all of these products gives us

$$z^{n-1}(-r_1 - r_2 - r_3 - \cdots - r_n)$$

as the z^{n-1} term in the expansion. However, the product $(z-r_1)(z-r_2)(z-r_3)\cdots(z-r_n)$ equals z^n-1, and there is no z^{n-1} term in the polynomial z^n-1. Therefore, the coefficient of z^{n-1} in the expansion of $(z-r_1)(z-r_2)(z-r_3)\cdots(z-r_n)$ must be 0. We just found that this coefficient is the opposite of the sum of the roots, so the sum of the roots is 0. □

Sidenote:

It's not just the sum of the roots that we can quickly determine from the coefficients of a polynomial. Consider the expansion of the product

$$(z-r_1)(z-r_2)(z-r_3)\cdots(z-r_n).$$

The highest degree term is z^n. As just explained, the coefficient of the z^{n-1} term in the expansion is $-(r_1+r_2+r_3+\cdots+r_n)$. But what about the z^{n-2} term in the expansion?

To form a z^{n-2} term, we can take $-r_1$ and $-r_2$ from the first two factors, and z from the other $n-2$ factors, which gives us $r_1r_2z^{n-2}$. We can do similarly to form every possible term $r_ir_jz^{n-2}$ with $i\neq j$, so the z^{n-2} term in the expansion is

$$(r_1r_2+r_1r_3+r_1r_4+\cdots+r_{n-1}r_n)z^{n-2}.$$

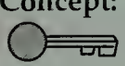
Therefore, the coefficient of z^{n-2} is the sum of all possible pairs of roots.

The coefficient of z^{n-1} is $(-1)^1$ times the sum of the roots. The coefficient of z^{n-2} is $(-1)^2$ times the sum of all possible products of two of the roots. Continuing in this vein, we can show that the coefficient of z^{n-3} is $(-1)^3$ times the sum of all possible products of three of the roots, z^{n-4} is $(-1)^4$ times the sum of all possible products of four of the roots, and so on.

These relationships are often referred to as **Vieta's formulas**, and they are covered extensively in Art of Problem Solving's *Intermediate Algebra* book.

Problem 7.31: Evaluate the sum $\cos\dfrac{2\pi}{2n+1}+\cos\dfrac{2\cdot2\pi}{2n+1}+\cos\dfrac{3\cdot2\pi}{2n+1}+\cdots+\cos\dfrac{n\cdot2\pi}{2n+1}$, where n is a positive integer.

Solution for Problem 7.31: The sum consists of terms of the form $\cos\frac{k\cdot2\pi}{2n+1}$. These are the real parts of some of the $(2n+1)^{\text{th}}$ roots of unity.

Concept: 

Problems involving the real or imaginary parts of roots of unity can often be solved using complex numbers, so keep an eye out for cosines or sines that are the real or imaginary parts of roots of unity.

The given sum has the real parts of only n of the $(2n+1)^{\text{th}}$ roots of unity. It's not clear how to handle such a sum, so we first look at a simpler problem.

Concept:

We can often get hints for how to tackle a hard problem by first tackling a simpler version of the problem.

One simpler version of the problem is to consider the real parts of all the m^{th} roots of unity:

$$\cos\frac{2\pi}{m}+\cos\frac{2\cdot2\pi}{m}+\cos\frac{3\cdot2\pi}{m}+\cdots+\cos\frac{m\cdot2\pi}{m}.$$

While we haven't tackled a sum like this, we have tackled a closely related sum—the sum of the n^{th} roots of unity themselves. We just showed that the sum of the m^{th} roots of unity is 0. Therefore, the sum of the real parts of the

m^{th} roots of unity is also 0 (and so is the sum of the imaginary parts). This gives us

$$\cos\frac{2\pi}{m} + \cos\frac{2\cdot 2\pi}{m} + \cos\frac{3\cdot 2\pi}{m} + \cdots + \cos\frac{m\cdot 2\pi}{m} = 0.$$

How does this help us with our original problem? Letting $m = 2n+1$ in the equation above produces

$$\cos\frac{2\pi}{2n+1} + \cos\frac{2\cdot 2\pi}{2n+1} + \cos\frac{3\cdot 2\pi}{2n+1} + \cdots + \cos\frac{(2n+1)\cdot 2\pi}{2n+1} = 0.$$

This contains all the terms in the desired sum, plus a bunch more terms:

$$\underbrace{\cos\frac{2\pi}{2n+1} + \cos\frac{2\cdot 2\pi}{2n+1} + \cdots + \cos\frac{n\cdot 2\pi}{2n+1}}_{\text{Desired sum}} + \underbrace{\cos\frac{(n+1)\cdot 2\pi}{2n+1} + \cos\frac{(n+2)\cdot 2\pi}{2n+1} + \cdots + \cos\frac{(2n+1)\cdot 2\pi}{2n+1}}_{\text{Extra terms}} = 0. \quad (7.1)$$

We do at least know how to handle one of the extra terms: the last one simplifies to $\cos 2\pi$, which is 1. But what about the rest?

Here are two methods for tackling the rest of terms and finishing the problem:

Method 1: Pair up the angles in the sum. We can often handle complicated sums (or products) of trig expressions by exploiting relationships among the angles in the terms. Here, we can pair each angle in the desired sum with an angle in the "extra terms" such that the sum of the angles is 2π. For example, $\frac{2\pi}{2n+1} + \frac{(2n)\cdot 2\pi}{2n+1} = \frac{(2n+1)\cdot 2\pi}{2n+1} = 2\pi$. Since $\cos(2\pi - \theta) = \cos\theta$, we have

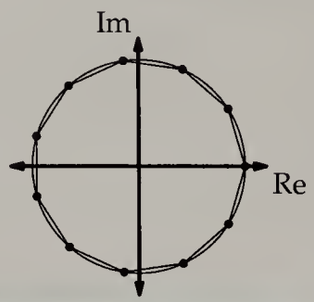
$$\cos\frac{(n+1)\cdot 2\pi}{2n+1} + \cos\frac{(n+2)\cdot 2\pi}{2n+1} + \cdots + \cos\frac{(2n)\cdot 2\pi}{2n+1} = \cos\frac{(n)\cdot 2\pi}{2n+1} + \cos\frac{(n-1)\cdot 2\pi}{2n+1} + \cdots + \cos\frac{1\cdot 2\pi}{2n+1}.$$

Hence, we can now write Equation (7.1) as

$$2\left(\cos\frac{2\pi}{2n+1} + \cos\frac{2\cdot 2\pi}{2n+1} + \cdots + \cos\frac{n\cdot 2\pi}{2n+1}\right) + 1 = 0,$$

from which we see that the desired sum equals $-\frac{1}{2}$.

Method 2: Geometry. We can visualize the $(2n+1)^{\text{th}}$ roots of unity as points equally spaced on the circumference of the unit circle, starting with the root at 1, as shown at right. The picture tells the story. The roots above the x-axis correspond to the terms in the desired sum. The roots on and below the x-axis are the "extra terms." The roots below the x-axis are the reflections of the "desired" roots over the x-axis, so the real part of each of these roots equals the real part of one of the "desired" roots. Therefore, the sum of the real parts of all the roots is twice the desired sum plus 1. Since the sum of the real parts of all of the roots must equal 0, we conclude that the desired sum is $-\frac{1}{2}$. □



Problem 7.32: The equation

$$x^{10} + (13x-1)^{10} = 0$$

has the 10 complex roots $r_1, \overline{r_1}, r_2, \overline{r_2}, r_3, \overline{r_3}, r_4, \overline{r_4}, r_5, \overline{r_5}$. Find the value of

$$\frac{1}{r_1\overline{r_1}} + \frac{1}{r_2\overline{r_2}} + \frac{1}{r_3\overline{r_3}} + \frac{1}{r_4\overline{r_4}} + \frac{1}{r_5\overline{r_5}}.$$

(Source: AIME)

Solution for Problem 7.32: We sure don't want to multiply out $(13x-1)^{10}$. Perhaps we can rewrite the equation in a more convenient form. Subtracting $(13x-1)^{10}$ from both sides gives

$$x^{10} = -(13x-1)^{10}.$$

This is close to an equation we know how to solve. We know what to do with $x^{10} = 1$, and what to do with $x^{10} = c$ for some constant c. While we're not so sure what to do with x^{10} equal to an expression, we can rewrite this equation as the tenth power of an expression equal to a constant by dividing both sides of $x^{10} = -(13x-1)^{10}$ by $(13x-1)^{10}$, which gives

$$\left(\frac{x}{13x-1}\right)^{10} = -1.$$

(Note that we can take this step without worrying about dividing by 0 because $x = 1/13$ is clearly not a solution to $x^{10} = -(13x-1)^{10}$, since $x = 1/13$ makes the right side 0 and the left side nonzero.)

Now, we have an equation that's a lot more familiar.

Concept: A great many equations can be solved by manipulating the equation into a form you already know how to tackle.

To make this equation *exactly* like an equation we have solved before, we can let $t = \frac{x}{13x-1}$, so our equation is simply $t^{10} = -1$. We can now use Problem 7.22 as a guide to find t. We have $t^{10} = -1 = e^{\pi i}$, so the 10 solutions for t are $e^{\pi i/10}$ times each of the tenth roots of unity. Since we can express each tenth root of unity as $e^{2\pi ki/10}$ for some integer k with $0 \le k < 10$, each solution for t has the form $t = e^{\pi i/10} \cdot e^{2\pi ki/10} = e^{(2k+1)\pi i/10}$ for some integer k.

We now must find x in terms of t. Multiplying both sides of $t = \frac{x}{13x-1}$ by $13x-1$ gives $13tx - t = x$, so $x = \frac{t}{13t-1}$. Therefore,

$$x = \frac{t}{13t-1} = \frac{e^{(2k+1)\pi i/10}}{13e^{(2k+1)\pi i/10} - 1}$$

is a solution of the given equation for any integer k.

That's not a pretty expression, but we aren't asked for the solutions of the original equation—we must evaluate an expression in terms of these solutions. Hopefully, once we plug these solutions into the desired expression, something convenient will happen. Let's try the first term in the desired expression, $\frac{1}{r_1\overline{r_1}}$. If $r_1 = \frac{t}{13t-1}$ is one solution, then

$$\begin{aligned}\frac{1}{r_1\overline{r_1}} = \frac{1}{r_1}\cdot\overline{\left(\frac{1}{r_1}\right)} &= \left(\frac{13t-1}{t}\right)\overline{\left(\frac{13t-1}{t}\right)} \\ &= \left(\frac{13t-1}{t}\right)\left(\frac{13\bar{t}-1}{\bar{t}}\right) \\ &= \frac{(13t-1)(13\bar{t}-1)}{t\bar{t}} \\ &= \frac{169t\bar{t} - 13t - 13\bar{t} + 1}{t\bar{t}}.\end{aligned}$$

Before we substitute for t, we notice that $t\bar{t} = |t|^2 = 1$ because $t = e^{(2k+1)\pi i/10}$ and $|e^{i\theta}| = 1$ for any angle θ. Therefore, we have

$$\frac{1}{r_1\overline{r_1}} = \frac{169t\bar{t} - 13t - 13\bar{t} + 1}{t\bar{t}} = \frac{169 - 13(t+\bar{t}) + 1}{1} = 170 - 13(t+\bar{t}).$$

Since $\bar{t} = \overline{e^{(2k+1)\pi i/10}} = e^{-(2k+1)\pi i/10}$, we have

$$\frac{1}{r_1\overline{r_1}} = 170 - 13\left(e^{(2k+1)\pi i/10} + e^{-(2k+1)\pi i/10}\right).$$

We still have two exponentials in this expression, but their exponents are opposites. We've seen this before, back in Problem 7.16, where we showed that $\cos\theta = (e^{i\theta} + e^{-i\theta})/2$. So, we have $e^{i\theta} + e^{-i\theta} = 2\cos\theta$. Applying this to our expression above gives us

$$\frac{1}{r_1\overline{r_1}} = 170 - 13\left(e^{(2k+1)\pi i/10} + e^{-(2k+1)\pi i/10}\right) = 170 - 26\cos\left(\frac{(2k+1)\pi}{10}\right).$$

This expression has 5 different possible values, one for each of $k = 0, 1, 2, 3$, and 4. When $k = 5$, we have $\cos\left(\frac{(2k+1)\pi}{10}\right) = \cos\frac{11\pi}{10}$, which equals $\cos\frac{9\pi}{10}$, so our expression for $\frac{1}{r_1\overline{r_1}}$ is the same for $k = 5$ as it is for $k = 4$. Similarly, every integer k gives us a value of our expression $\frac{1}{r_1\overline{r_1}}$ that equals one of the 5 values we find for $k = 0, 1, 2, 3$, or 4.

Since there is a solution to $x^{10} + (13x - 1)^{10} = 0$ of the form $x = e^{(2k+1)\pi i/10}/(13e^{(2k+1)\pi i/10} - 1)$ for each of the tenth roots of unity, we know that among the five terms of the form $\frac{1}{r_1\overline{r_1}}$, each of the 5 possible values of $170 - 26\cos\left(\frac{(2k+1)\pi}{10}\right)$ occurs. Therefore, our desired sum equals

$$5(170) - 26\left(\cos\frac{\pi}{10} + \cos\frac{3\pi}{10} + \cos\frac{5\pi}{10} + \cos\frac{7\pi}{10} + \cos\frac{9\pi}{10}\right).$$

We're almost there! We have $\cos\frac{5\pi}{10} = \cos\frac{\pi}{2} = 0$. The other two cosines can be grouped in pairs whose angles sum to π, which reminds us that $\cos x = -\cos(\pi - x)$. So, we have $\cos\frac{\pi}{10} + \cos\frac{9\pi}{10} = \cos\frac{3\pi}{10} + \cos\frac{7\pi}{10} = 0$. Therefore, the sum of cosines in our expression for the desired sum is 0, so the sum equals $5(170) = 850$. □

Exercises

7.4.1

(a) Find the roots of $z^2 - z + 1$ in exponential form.

(b) Find the roots of $z^5 + 1$ in exponential form.

(c) Find a degree 4 polynomial whose roots are the nonreal roots you found in part (b).

(d) Find the roots of $x^6 - x^5 + x^4 - x^3 + x^2 - x + 1$. (You can leave your answers in exponential form.)

7.4.2 Let ω be a 13^{th} root of unity. Evaluate $(1 - \omega)(1 - \omega^2)\cdots(1 - \omega^{12})$.

7.4.3 Let z be such that $z^7 = 1$ and $z \neq 1$. Compute the numerical value of

$$z^{10} + \frac{1}{z^{10}} + z^{30} + \frac{1}{z^{30}} + z^{50} + \frac{1}{z^{50}}.$$

(Source: NYSML)

7.4.4 The nonzero complex numbers a, b, and c satisfy $\frac{a}{b} = \frac{b}{c} = \frac{c}{a}$. What are the possible values of $\frac{a+b-c}{a-b+c}$?

7.4.5 Factor $x^7 + x^4 + x^3 + x + 1$ into the product of two polynomials with integer coefficients and positive degrees.

7.4.6⋆ The polynomial

$$P(x) = (1 + x + x^2 + \cdots + x^{17})^2 - x^{17}$$

has 34 roots of the form $z_k = r_k[\cos(2\pi\alpha_k) + i\sin(2\pi\alpha_k)]$, $k = 1, 2, \ldots, 34$, with $0 < \alpha_1 \leq \alpha_2 \leq \alpha_3 \leq \cdots \leq \alpha_{34} < 1$ and $r_k > 0$. Find $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5$. *(Source: AIME)* **Hints:** 209, 58

7.5 Summary

The form $x + yi$ is the **rectangular form** of a complex number. A complex number can also be written in the **polar form** $r(\cos\theta + i\sin\theta)$ or the **exponential form** $re^{i\theta}$. We typically have $r \geq 0$ in polar and exponential form, so that r is the **magnitude** of the complex number. The θ in polar and exponential form is called **the argument** of the complex number. The rectangular form of a complex number is related to the polar and exponential forms by $x = r\cos\theta$, $y = r\sin\theta$, so $r^2 = x^2 + y^2$ and $\tan\theta = \frac{y}{x}$ (when $\cos\theta$ is nonzero).

Important: **De Moivre's Theorem** states that

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$$

for any θ and any positive integer n.

Important: If $z = re^{i\theta}$, then $\bar{z} = re^{-i\theta}$.

Important: For any angle θ, we have

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2},$$
$$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

Important: The n roots of $f(x) = x^n - 1$ are

$$e^0, e^{2\pi i/n}, e^{4\pi i/n}, e^{6\pi i/n}, \ldots, e^{2(n-1)\pi i/n}.$$

These are also referred to as the **n^{th} roots of unity**, since these are the n complex numbers whose n^{th} power is 1.

Important: An n^{th} root of unity, z, is called a **primitive n^{th} root of unity** if $z^n = 1$ and $z^k \neq 1$ for $k = 1, 2, 3, \ldots, n-1$. If $\gcd(k, n) = 1$, then $\omega = e^{2\pi ik/n}$ is a primitive n^{th} root of unity, and $1, \omega, \omega^2, \omega^3, \ldots, \omega^{n-1}$ are all of the n^{th} roots of unity.

Things To Watch Out For!

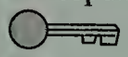
We always use radians in exponential notation, never degrees.

Problem Solving Strategies

Concepts:

- Investigating examples can reveal patterns that help conjecture and prove general results.
- When working with a difficult expression, think about where else you have seen similar expressions.

Concepts:



- We can often get hints for how to tackle a hard problem by first tackling a simpler version of the problem.
- A great many equations can be solved by manipulating the equation into a form you already know how to tackle.
- De Moivre's Theorem can help us tackle problems involving $\cos n\theta$ or $\sin n\theta$, where n is a positive integer greater than 1.
- Exponential form sometimes offers an easy way to prove facts involving complex number multiplication.
- The exponential form of complex numbers is a helpful way to remember trigonometric identities.
- Many problems involving polynomials of the form $x^n + x^{n-1} + x^{n-2} + \cdots + 1$ can be solved by writing the polynomial in the form $\frac{x^{n+1}-1}{x-1}$, and using our knowledge of the roots of unity.
- Problems involving the real or imaginary parts of roots of unity can often be solved using complex numbers, so keep an eye out for cosines or sines that are the real or imaginary parts of roots of unity.

REVIEW PROBLEMS

7.33 Write each of the following in polar form:

(a) -9
(b) $\sqrt{6} + \sqrt{2}i$
(c) $8 + 8i$
(d) $-\frac{1}{2} - \frac{1}{2}i$

7.34 Write each of the following in rectangular form:

(a) $4 \operatorname{cis} 120°$
(b) $\frac{1}{2} \operatorname{cis} 315°$
(c) $8 \operatorname{cis} \frac{\pi}{6}$
(d) $20 \operatorname{cis} 97\pi$

7.35 For what values of θ does $\operatorname{cis}\theta = \operatorname{cis} 2\theta$?

7.36 Determine all integer values of θ with $0° \le \theta \le 90°$ for which $(\cos\theta + i\sin\theta)^{75}$ is a real number. *(Source: ARML)*

7.37 Express each of the following in exponential form:

(a) $24 - 24i$
(b) $\frac{4\sqrt{3}}{3} + 4i$

7.38 Express each of the following in rectangular form:

(a) $e^{3\pi i/2}$
(b) $6e^{7\pi i/6}$

7.39 Use the fact that $e^{-i\theta} = \frac{1}{e^{i\theta}}$ to reproduce the identities for $\sin(-\theta)$ and $\cos(-\theta)$.

7.40 Describe all complex numbers z such that $\left|\frac{z}{\overline{z}} + \frac{\overline{z}}{z}\right| = \sqrt{3}$.

7.41 Evaluate each of the following:

(a) $\left(\frac{\sqrt{3}}{4} - \frac{1}{4}i\right)^8$ (b) $\frac{(-3\sqrt{3} - 3i)^8}{(1 + \sqrt{3}i)^6}$

7.42

(a) Express $\sin 5\theta$ in terms of $\sin\theta$.

(b) Show that $\sin 4\theta$ cannot be expressed as a function of $\sin\theta$.

7.43 If $w \cdot z$ is a nonzero real number and $|w| = |z|$, then which of the following is true:

(a) $w + z$ is real.

(b) $w + z$ is either real or pure imaginary.

(c) It is possible that $w + z$ is neither real nor pure imaginary.

7.44 Find all solutions to $z^3 = -4\sqrt{2} + 4\sqrt{2}i$ in exponential form.

7.45 Let $z = e^{i\theta}$, where $z \neq 1$. Show that $\frac{1+z}{1-z} = i\cot\frac{\theta}{2}$.

7.46

(a) Find the product of the n^{th} roots of unity.

(b) Find the product of the primitive n^{th} roots of unity.

7.47 Let n be a positive integer. How are the roots of $x^{2n+1} + 1$ related to the roots of $x^{2n+1} - 1$

(a) algebraically?

(b) geometrically, in the complex plane?

7.48 Let w be a nonzero complex number. Suppose the solutions to the equation $z^n = w$ are graphed in the complex plane. Prove that the resulting points are vertices of a regular n-gon.

7.49 Let $x_n + iy_n = (1 + i\sqrt{3})^n$, where x_n and y_n are real and n is a positive integer. If $x_{19}y_{91} + x_{91}y_{19} = 2^k\sqrt{3}$, compute k. *(Source: ARML)*

7.50 Show that $\overline{e^z} = e^{\bar{z}}$ for all complex numbers z.

7.51 The **hyperbolic sine**, **hyperbolic cosine**, and **hyperbolic tangent** functions are defined and denoted as follows:

$$\sinh x = \frac{e^x - e^{-x}}{2}, \qquad \cosh x = \frac{e^x + e^{-x}}{2}, \qquad \tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

Prove the following identities:

(a) $\cosh^2 x - \sinh^2 x = 1$.

(b) $\sinh(x + y) = \sinh x \cosh y + \cosh x \sinh y$.

(c) $\tanh(x + y) = \dfrac{\tanh x + \tanh y}{1 + \tanh x \tanh y}$.

(d) $\sinh x + \sinh y = 2\sinh\frac{x+y}{2}\cosh\frac{x-y}{2}$.

7.52 If $x = \frac{-1+i\sqrt{3}}{2}$ and $y = \frac{-1-i\sqrt{3}}{2}$, then find all nonnegative integers n such that $x^n + y^n = -1$.

Challenge Problems

7.53 The equation $z^6 + z^3 + 1 = 0$ has one complex root with argument (angle) θ between $90°$ and $180°$ in the complex plane. Determine the degree measure of θ. *(Source: AIME)*

7.54 Let z be a root of $x^5 - 1 = 0$, with $z \neq 1$. Compute the value of

$$z^{15} + z^{16} + z^{17} + \cdots + z^{50}.$$

(Source: ARML)

7.55 Find all solutions to the equation $z^7 + 8z^4 - 2iz^3 - 16i = 0$. **Hints:** 39

7.56 Let $\theta = \arctan 2$. What quadrant does the angle 8θ lie in? (No calculators!) **Hints:** 192

7.57 Find all z such that $z^9 + z^6 + z^3 = -1$. **Hints:** 170, 283

7.58 Describe all integers n such that the polynomial $x^{2n} + 1 + (x+1)^{2n}$ is divisible by $x^2 + x + 1$. **Hints:** 254

7.59 Find all z such that $32z^5 + 16z^4 + 8z^3 + 4z^2 + 2z + 1 = 0$. **Hints:** 12

7.60 Let x and y be two k^{th} roots of unity. Prove that $(x+y)^k$ is real. *(Source: HMMT)* **Hints:** 38

7.61 If $\omega^{1997} = 1$ and $\omega \neq 1$, then evaluate

$$\frac{1}{1+\omega} + \frac{1}{1+\omega^2} + \cdots + \frac{1}{1+\omega^{1997}}.$$

(Source: Mandelbrot) **Hints:** 214

7.62 A sequence of complex numbers $z_0, z_1, z_2, \ldots$ satisfies the rule

$$z_{n+1} = \frac{iz_n}{\overline{z}_n}.$$

Suppose that $|z_0| = 1$ and $z_{2005} = 1$. How many possible values are there for z_0? *(Source: AMC 12)* **Hints:** 109, 70

7.63 The sequences $a_1, a_2, a_3, \ldots$ and $b_1, b_2, b_3, \ldots$ are defined by $a_1 = \alpha$, $b_1 = \beta$, and $a_{n+1} = \alpha a_n - \beta b_n$, $b_{n+1} = \beta a_n + \alpha b_n$ for all $n \geq 1$. How many pairs of real numbers (α, β) are there such that $a_{1997} = a_1$ and $b_{1997} = b_1$? **Hints:** 152, 132

7.64 Show that $\tan n\theta = \dfrac{\binom{n}{1}\tan\theta - \binom{n}{3}\tan^3\theta + \cdots}{1 - \binom{n}{2}\tan^2\theta + \cdots}$. **Hints:** 16

7.65 Given that z is a complex number such that $z + \frac{1}{z} = 2\cos 3°$, find the least integer that is greater than $z^{2000} + \frac{1}{z^{2000}}$. *(Source: AIME)* **Hints:** 257

7.66⋆ For a positive integer n, let $\Phi_n(x)$ denote the polynomial whose leading coefficient is 1, and whose roots are the primitive n^{th} roots of unity. These polynomials are called the **cyclotomic polynomials**.

(a) Show that $x^n - 1$ can be factored as $\Phi_{d_1}(x)\Phi_{d_2}(x)\Phi_{d_3}(x)\cdots\Phi_{d_k}(x)$, where $d_1, d_2, d_3, \ldots, d_k$ are all of the positive divisors of n. **Hints:** 141

(b) Find $\Phi_n(x)$ for $1 \leq n \leq 10$. **Hints:** 256

(c) Find $\Phi_p(x)$, where p is a prime.

(d) Show that if $n > 2$, then the degree of $\Phi_n(x)$ is even.

(e) Show that $\Phi_{2n}(x) = \Phi_n(-x)$ for all odd integers $n > 2$. **Hints:** 188

7.67★ Show that

$$\binom{n}{0} - \binom{n}{2} + \binom{n}{4} - \binom{n}{6} + \cdots = 2^{n/2} \cos \frac{n\pi}{4}$$

and

$$\binom{n}{1} - \binom{n}{3} + \binom{n}{5} - \binom{n}{7} + \cdots = 2^{n/2} \sin \frac{n\pi}{4}.$$

Hints: 272

7.68★ If $w = \cos 40^\circ + i \sin 40^\circ$, then express $|w + 2w^2 + 3w^3 + \cdots + 9w^9|^{-1}$ in the form $a \sin b$. *(Source: AHSME)* **Hints:** 165, 88

7.69★ The sequence of functions $\{T_n\}$ is defined by $T_0(x) = 1$, $T_1(x) = x$, and $T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x)$ for all $n \geq 2$. Show that $T_n(\cos\theta) = \cos n\theta$.

7.70★ The angles α, β, and γ satisfy $e^{i\alpha} + e^{i\beta} + e^{i\gamma} = 0$. Show that $e^{3i\alpha} = e^{3i\beta} = e^{3i\gamma}$. **Hints:** 190, 77

7.71★ Show that

$$\cos x + \cos y + \cos z + \cos(x + y + z) = 4 \cos\left(\frac{x+y}{2}\right) \cos\left(\frac{x+z}{2}\right) \cos\left(\frac{y+z}{2}\right)$$

and

$$\sin x + \sin y + \sin z - \sin(x + y + z) = 4 \sin\left(\frac{x+y}{2}\right) \sin\left(\frac{x+z}{2}\right) \sin\left(\frac{y+z}{2}\right).$$

Hints: 211, 130

7.72★ Find a cubic polynomial whose roots are $\cos 2\pi/7$, $\cos 4\pi/7$, and $\cos 6\pi/7$. **Hints:** 162, 25

7.73★ Let $f(x) = x^{2004} + 2x^{2003} + 3x^{2002} + \cdots + 2004x + 2005$ and $z = \cos\left(\frac{\pi}{1003}\right) + i \sin\left(\frac{\pi}{1003}\right)$. Express the product

$$f(z)f(z^2)\cdots f(z^{2005})$$

in the form a^b, where a and b are integers. **Hints:** 155, 73, 227

7.74★ Let v and w be distinct, randomly chosen roots of the equation $z^{1997} - 1 = 0$. Find the probability that $\sqrt{2 + \sqrt{3}} \leq |v + w|$. *(Source: AIME)* **Hints:** 139, 276

7.75★ Show that if n is a positive integer greater than 1, then

$$\sin\frac{\pi}{n} \sin\frac{2\pi}{n} \cdots \sin\frac{(n-1)\pi}{n} = \frac{n}{2^{n-1}}.$$

Hints: 37, 183, 142

Extra!

What is i^i?

De Moivre's Theorem allows us to raise nonreal numbers to integer powers, and our work with finding roots of unity gives us some idea how to handle raising nonreal numbers to rational powers. But what about raising a nonreal number to a nonreal power?

Let's start simpler, by raising a real number to a nonreal power. We know how to raise e to a nonreal power, but what is 2^i? The only number we know how to raise to the power i is e, so we write 2 as a power of e. By definition, the power that we must raise e to in order to get 2 is the base e logarithm of 2. We typically denote the base e logarithm as ln, so, for example, $\ln e^3 = 3$. Since we must raise e to the power $\ln 2$ to get 2, we can write 2 as $e^{\ln 2}$, so we have

$$2^i = \left(e^{\ln 2}\right)^i = e^{(\ln 2)i} = \cos(\ln 2) + i\sin(\ln 2).$$

So, now we know how to raise positive real numbers besides e to nonreal powers—express the base of the expression as a power of e.

We computed 2^i by writing 2 as $e^{\ln 2}$. So, to compute i^i, we try writing i as $e^{\ln i}$, which would allow us to write i^i as $(e^{\ln i})^i$. We don't yet know what $\ln i$ is, but we do know how to write i as a power of e. Specifically, we have $e^{\pi i/2} = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = i$, so we can choose to let $\ln i = \pi i/2$, and we have

$$i^i = \left(e^{\pi i/2}\right)^i = e^{(\pi i/2)\cdot i} = e^{-\pi/2}.$$

Weird! We raise i to the power i, and the result is a real number!

But wait! We also have $e^{5\pi i/2} = i$. So

$$i^i = \left(e^{5\pi i/2}\right)^i = e^{-5\pi/2}.$$

In fact we see that we could have gotten any number of the form $e^{-\pi/2+2k\pi}$ where k is an integer. There are a lot of possible values. What is going on here?

We'd like to define i^i such that it has only one value. Our work above seems to give many values of i^i, so one of our steps may involve using a "function" that isn't properly defined to give a unique output for each input. We've seen something like this before when we defined inverse trig functions. There, we had to restrict the domain of $\cos x$ in order to define an inverse. Let's see if something similar is going on here.

Remember we wrote $i = e^{\ln i}$. Raising e to a power is a function, since

$$e^{a+bi} = e^a \cdot e^{bi} = e^a(\cos b + i\sin b) = e^a\cos b + e^a i\sin b,$$

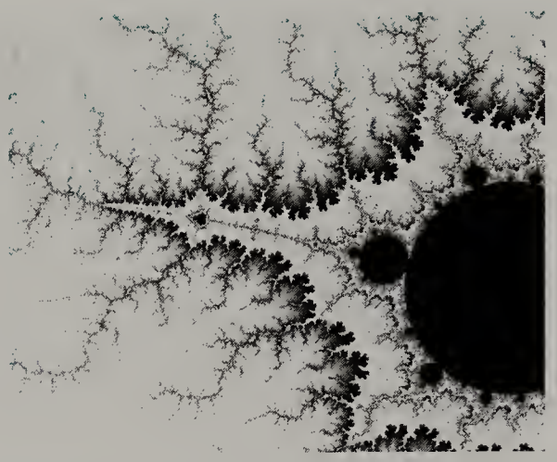
which definitely has only one possible value. But what about ln? Since we wrote $i = e^{\ln i}$, we're assuming that ln is an inverse to $f(x) = e^x$. In order for f to have an inverse, no two inputs can go to the same output. This is OK when x is real. When we allow nonreal inputs to $f(x) = e^x$, we have a different story, because $f(x) = f(x + 2\pi i)$.

Now we see the problem! In order to build the function ln we need to restrict the domain of f. (Remember this is exactly what we did in defining arccos, since cos itself did not have an inverse.) One of the most common choices for restricting the domain of f is forcing the imaginary part of the input to be in $(-\pi, \pi]$.

If we make this choice, then

$$i^i = e^{i\ln i} = e^{i(\pi i/2)} = e^{-\pi/2}.$$

As you might guess, this is far from the end of the story. Defining ln in this way introduces some unsettling consequences. For example, compare $\ln e^{\pi i}$ to $\ln(e^{1.0000001\pi i})$. The first equals πi. We'd expect the second to be quite close to πi, but, no, it's actually $-0.9999999\pi i$. Naturally, mathematicians have an answer for dealing with this oddity, too. Perhaps you'll study it in college when you study complex analysis or topology.



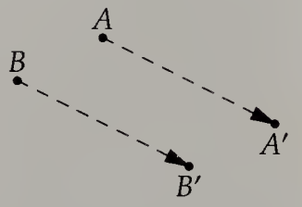
The shortest path between two truths in the real domain passes through the complex domain. – Jacques Hadamard

CHAPTER 8
Geometry of Complex Numbers

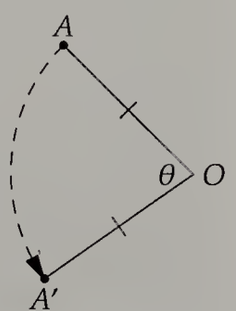
Many mathematical tools can be applied to a wide range of problems, including problems that seem very unrelated to each other. This chapter exhibits various deep relationships between two seemingly unrelated areas of mathematics: complex numbers and geometry. As you work through the following problems, keep in mind that many significant advances in mathematics have come from the realization that two seemingly different fields of mathematics are strongly related, or even the same.

8.1 Transformations of the Complex Plane

We start our investigation of the geometry of complex numbers with geometric transformations.

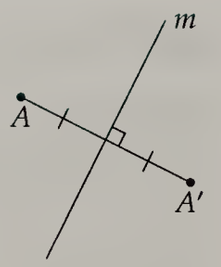


- **Translation**. When we apply a translation to a figure, we slide it a specified distance in a given direction. For example, points A' and B' at right are the result of translating A and B, respectively, in the direction and by the distance suggested by the arrows. So, we say that A' is the "image" of A under the translation shown at right, or that the translation "maps" A to A'. (We use the terms "image" and "maps" in this manner with all transformations.)

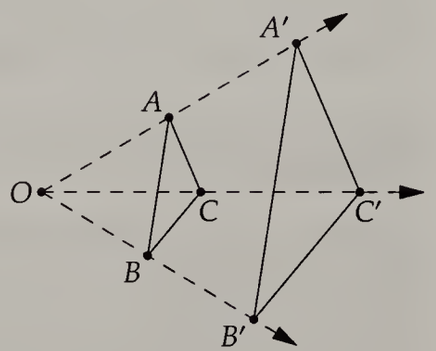


- **Rotation**. When we rotate a figure, we spin it by some angle about some point. For example, in the diagram at right, the image of point A under a rotation of angle θ counterclockwise about point O is the point A' such that $OA = OA'$ and $\angle AOA' = \theta$, as shown at right. If θ is not an integer multiple of π, then there are two points that satisfy this description of A', which is why we must specify the **orientation** of the rotation as clockwise or counterclockwise. We call point O the **center** of this rotation, and θ is the **angle** of rotation.

- **Reflection**. The image of point A upon reflection over line m is the point A' such that m is the perpendicular bisector of $\overline{AA'}$. In other words, if we fold our paper along line m, points A and A' will coincide. If point A is on m, then it is its own reflection. We call m the **axis of reflection**.



- **Dilation**. At right is an example of dilation. Point O is the **center of dilation**. The image of A is the point A' on $\overrightarrow{OA}$ such that $OA'/OA = 2$. Similarly, the image of point B is the point B' on $\overrightarrow{OB}$ such that $OB'/OB = 2$, and the image of point C is the point C' on $\overrightarrow{OC}$ such that $OC'/OC = 2$. The image of the center of dilation is itself. We can think of the dilation as "stretching" $\triangle ABC$ away from O, and we say that $\triangle A'B'C'$ is a dilation of $\triangle ABC$ about O.

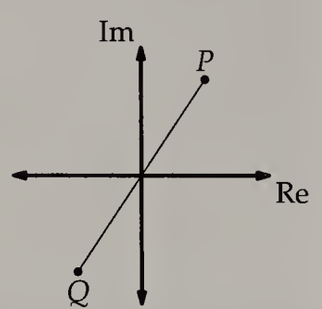


In general, the image of a point P upon dilation with **scale factor** k and center O is the point P' on $\overrightarrow{OP}$ such that $OP' = k(OP)$ (when k is positive). The scale factor need not be positive. If the scale factor k is negative, then the image of P is the point P' on the ray from O *going in the opposite direction from* P such that $OP' = |k|OP$.

We say that translations, rotations, and reflections are **isometries**, which means that they preserve distance. We can use this fact to show that a segment and its image under an isometry have the same length, and an angle and its image under an isometry have the same measure. A dilation is not an isometry, because it does not preserve length. It does, however, preserve angle measure, so, for example, $\angle ABC = \angle A'B'C'$ in the dilation shown above.

We sometimes wish to speak about geometric figures in the complex plane that are more complicated than points, such as segments or triangles. Doing so with the notation of complex numbers could be confusing. For example, "wz" typically refers to the product of complex numbers w and z, not the segment from w to z in the complex plane. So, we usually stick to the notation of geometry when referring to such geometric objects in the complex plane. By convention, we use capital letters to denote points and the corresponding lowercase letters to denote corresponding complex numbers.

For example, in the diagram at right, point P corresponds to the complex number p and point Q corresponds to the complex number q. We refer to the segment that connects p and q as $\overline{PQ}$. While we usually use capital letters for any geometric object besides a point, we will sometimes refer to points in the complex plane by the complex numbers to which they correspond. So, we could describe the situation depicted at right as either "p is a 180° rotation of q about the origin" or "P is a 180° rotation of Q about the origin."



Problems

Problem 8.1:

(a) What transformation maps any complex number z to the complex number $z + 2 - 3i$?

(b) What transformation maps any complex number z to $z + w$, where w is a constant complex number?

Problem 8.2:

(a) If $z = \cos 12° + i \sin 12° + \cos 48° + i \sin 48°$, then what is the argument of z?

(b) Find $\text{Im}\left((\cos 12° + i \sin 12° + \cos 48° + i \sin 48°)^6\right)$.

Problem 8.3: If ω and $\omega - i$ are both n^{th} roots of unity, then what are the possible values of n?

Problem 8.4:

(a) What transformation maps z to $2z$?

(b) What transformation maps z to cz, where c is a positive real constant?

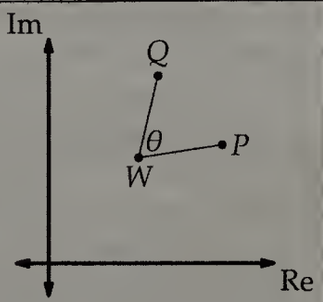
Problem 8.5:

(a) How are z and iz related to each other in the complex plane?

(b) Generalize your result from part (a) by finding a relationship between z and $e^{i\alpha}z$ in the complex plane.

Problem 8.6: Describe in words the transformation that maps z to $re^{i\theta} \cdot z$, where r is a positive real constant and θ is a real constant.

Problem 8.7: Let q be the result of rotating p an angle θ counterclockwise about w in the complex plane. In this problem, we find q in terms of θ, p, and w.



(a) Let points P, Q, and W correspond to complex numbers p, q, and w, respectively. Consider the translation that maps W to the origin, O. Let this translation map P to P' and Q to Q'. Find the complex numbers that correspond to P' and Q' in terms of p, q, and w.

(b) What is $\angle P'OQ'$?

(c) Why are P' and Q' equidistant from the origin?

(d) Find q in terms of p, w, and θ.

Problem 8.8: A pirate buries his treasure on a deserted island with an oak tree, a pine tree, and a gallows. To choose the burial location, he starts at the gallows and walks to the oak tree, counting his steps. When he reaches the oak tree, he turns right 90 degrees, walks the same number of steps as from the gallows to the oak tree, and places a spike in the ground.

He returns to the gallows, and then walks to the pine tree, again counting his steps. When he reaches the pine tree, he turns left 90 degrees, walks the same number of steps as from the gallows to the pine tree, and places another spike in the ground. He then buries the treasure at the midpoint of the segment with the spikes as endpoints, and then removes the spikes from the ground.

Years later, the pirate returns to the island to get his treasure. Unfortunately, the gallows are gone, and there is no sign of where it had been. Fortunately, the oak tree and pine tree are still there. In this problem, we help the pirate find his treasure. We do so by putting the problem on the complex plane.

(a) Let g correspond to the gallows, p to the pine, and k to the oak. Find complex numbers in terms of g, p, and k that correspond to each spike and to the treasure.

(b) When we place a geometry problem in the complex plane, we can choose where to place the origin. Suppose we choose the midpoint of the segment connecting the two trees as our origin. Then, how are p and k related? Why is this choice of origin so convenient?

(c) Where is the treasure?

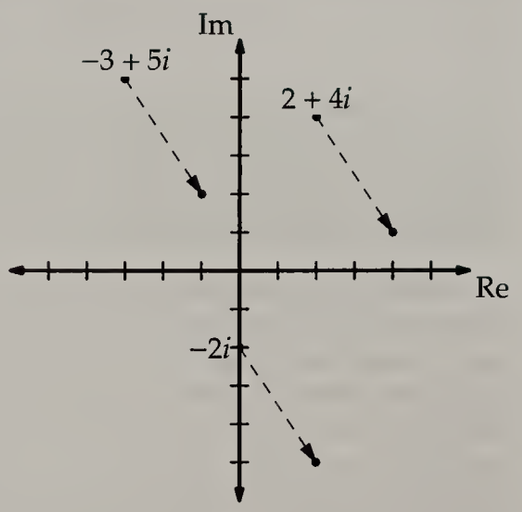
Problem 8.1:

(a) What transformation maps any complex number z to the complex number $z + 2 - 3i$?

(b) What transformation maps any complex number z to $z + w$, where w is a constant complex number?

Solution for Problem 8.1:

(a) We investigate by trying a few values of z. At right, we have plotted $2+4i$, $-3+5i$, and $-2i$, and the results when we add $2-3i$ to each. In each case, the result of the addition is 2 units to the right and 3 units below the original complex number. Writing z as $a+bi$ makes it clear why $z+2-3i$ is a translation of z:



$$z+2-3i = a+bi+2-3i = (a+2)+(b-3)i.$$

The effect of adding $2-3i$ to z is to add 2 to the real part (move 2 units right) and subtract 3 from the imaginary part (move 3 units down). Therefore, the result is a translation of z by 2 units right and 3 units down.

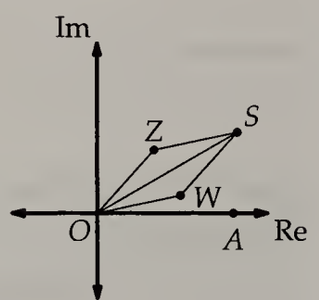
(b) From our work in part (a), we aren't surprised to find that the transformation that maps z to $z+w$ is a translation (even if $w=0$, which gives us the "do nothing" translation). If we let $w=c+di$, we see that $z+w$ adds c to the real part of z (moves c units horizontally) and adds d to the imaginary part of z (moves d units vertically). Therefore, $z+w$ is a translation of z by c units horizontally and d units vertically.

□

Problem 8.2: Find $\text{Im}\left((\cos 12^\circ + i\sin 12^\circ + \cos 48^\circ + i\sin 48^\circ)^6\right)$.

Solution for Problem 8.2: We could start by raising that whole expression to the sixth power, but that looks difficult. The angles are conveniently related, so we might try trig identities. This approach will work, but thinking about the problem geometrically gives a much faster solution.

Letting $w=\cos 12^\circ + i\sin 12^\circ$ and $z=\cos 48^\circ + i\sin 48^\circ$, our problem is to determine $\text{Im}\left((w+z)^6\right)$. We know how to find $w+z$ in the complex plane when we have w and z. Specifically, we let points W, Z, and S in the complex plane correspond to w, z, and $w+z$, respectively, as shown at right. As we saw in Problem 8.1, adding a constant complex number w corresponds to a translation in the complex plane. Adding w to 0 and to z give w and $w+z$, respectively. So, the same translation that maps O to W also maps Z to S. Therefore, $OW = ZS$ and $\overline{OW} \parallel \overline{ZS}$, which means that $OWSZ$ is a parallelogram. (We also saw this when we introduced graphing on the complex plane in Section 6.2.) Since $|w| = |z|$, we have $OZ = OW$, so this parallelogram is also a rhombus.



Since $OWSZ$ is a rhombus, diagonal $\overline{OS}$ bisects $\angle ZOW$. From the arguments of W and Z, we know that $\angle WOA = 12^\circ$ and $\angle ZOA = 48^\circ$, so $\angle WOZ = 36^\circ$ and $\angle SOW = \frac{1}{2}\angle WOZ = 18^\circ$. This means that $\angle SOA = 30^\circ$, which is an angle we like. Since $\angle SOA = 30^\circ$, the argument of $w+z$ is 30°, which means we have $w+z = re^{\pi i/6}$ for some positive constant r. Therefore, $(w+z)^6 = \left(re^{\pi i/6}\right)^6 = r^6 e^{\pi i} = r^6(-1) = -r^6$. We don't know exactly what the real part of this number is, but we know its imaginary part is 0, so $\text{Im}((w+z)^6) = 0$. □

Our solution to Problem 8.2 looks a lot longer than it really is. Once you're comfortable with the geometry of complex numbers, the observation that $w+z$ has argument 30° is almost immediate, which makes determining that $\text{Im}((w+z)^6) = 0$ very easy.

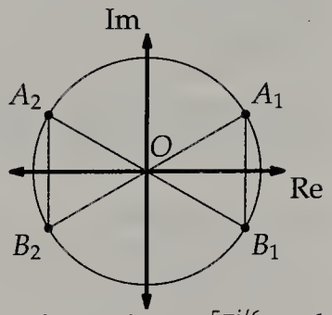
Let's take another look at a problem that is simplified by thinking of adding complex numbers as a translation.

Problem 8.3: If ω and $\omega - i$ are both n^{th} roots of unity, then what are the possible values of n?

Solution for Problem 8.3: If we start algebraically, we have $\omega^n = (\omega - i)^n = 1$. Um, then what? We sure don't want to expand $(\omega - i)^n$. We might write ω and $\omega - i$ in polar form and try to bash away with sines and cosines (you can take a swing at that on your own), but before we do that, let's think about the problem geometrically.

We let points A and B in the complex plane correspond to ω and $\omega - i$, respectively. Since ω and $\omega - i$ are both roots of unity, both are on the unit circle. The point $\omega - i$ is a one-unit downward translation of ω, so $\overline{AB}$ is vertical with $AB = 1$. Moreover, since A and B are on the unit circle, we have $OA = OB = 1$, which means that $\triangle OAB$ is equilateral. Therefore, $\angle AOB = 60°$.

Since $\overline{AB}$ is vertical and no two points on the unit circle above the real axis are on the same vertical line, we know that A and B cannot both be above the real axis. Similarly, they cannot both be below it and neither can be on it, so one is above the real axis and the other is below it. Since B is below A, we know that A is above the real axis and B is below it. The top half of the unit circle is the reflection of the bottom half over the real axis, so the real axis must bisect $\angle AOB$. This tells us that $\overline{AO}$ makes a 30° angle with the real axis, limiting A and B to the two possibilities shown in the diagram, with points A_1 and B_1 forming one possible pair and points A_2 and B_2 forming the other possible pair.



For the first possibility we have $A_1 = e^{\pi i/6}$ and $B_1 = e^{-\pi i/6}$, and for the second we have $A_2 = e^{5\pi i/6}$ and $B_2 = e^{-5\pi i/6}$. Now, we're ready to find n.

If $\omega = e^{\pi i/6}$ is an n^{th} root of unity, we must have $\left(e^{\pi i/6}\right)^n = 1$, so $e^{\pi i n/6} = 1$. This means that $\pi i n/6$ must be an integer multiple of 2π, so $\pi i n/6 = 2\pi i k$ for some integer k. Solving for n gives $n = 12k$, which means that n is a multiple of 12. Similarly, if $\omega = e^{5\pi i/6}$, then we have $\left(e^{5\pi i/6}\right)^n = 1$. This gives us $5\pi i n/6 = 2\pi i k$ for some integer k, from which we find $5n = 12k$. Once again, this means that n is a multiple of 12.

We must still confirm in both cases that $\omega - i$ is an n^{th} root of unity when n is a multiple of 12. When $\omega = e^{\pi i/6}$, we have $\omega - i = e^{-\pi i/6}$ and

$$(\omega - i)^n = \left(e^{-\pi i/6}\right)^n = \left(e^{-\pi i/6}\right)^{12k} = e^{-2\pi i k} = 1,$$

as desired. Similarly, when $\omega = e^{5\pi i/6}$, we have $\omega - i = e^{-5\pi i/6}$ and

$$(\omega - i)^n = \left(e^{-5\pi i/6}\right)^n = \left(e^{-5\pi i/6}\right)^{12k} = e^{-10\pi i k} = 1.$$

(We could also have noted that in both cases, we have $\omega - i = \overline{\omega}$, and we showed in Problem 7.21 that if ω is a root of unity, then so is $\overline{\omega}$.)

We can also use geometry to visualize the fact that n must be a multiple of 12. Our options for A and B are all primitive 12^{th} roots of unity, so the first 12 powers of each form the vertices of a regular dodecagon in the complex plane, as shown at left below.

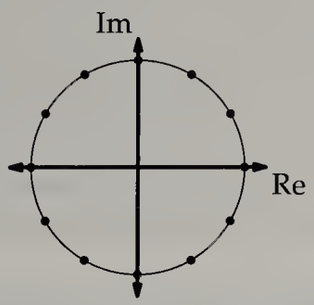


Figure 8.1: 12^{th} Roots of Unity

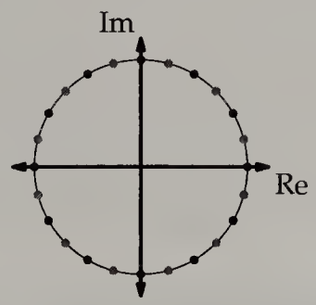


Figure 8.2: 24^{th} Roots of Unity

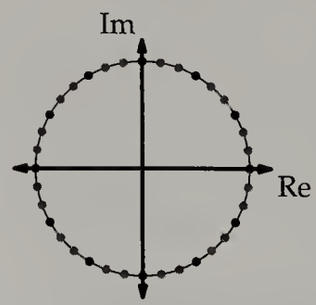


Figure 8.3: 36^{th} Roots of Unity

The diagrams above exhibit how the 12^{th} roots of unity are related to the 24^{th} and 36^{th} roots of unity. Similarly, whenever n is a multiple of 12, the n^{th} roots of unity include the 12^{th} roots of unity. Moreover, if a regular n-gon

with all its vertices on the unit circle includes these 12 roots of unity among its vertices, then n must be a multiple of 12, because there will be an equal number of vertices of the n-gon between each two consecutive 12^{th} roots of unity. □

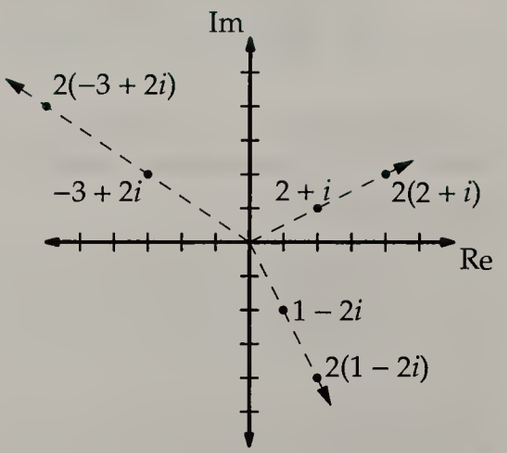
Problem 8.4:
(a) What transformation maps z to $2z$?
(b) What transformation maps z to cz, where c is a positive real constant?

Solution for Problem 8.4:

(a) We start by exploring a few examples. In the diagram at right, we have $2 + i$, $-3 + 2i$, and $1 - 2i$, and the results of multiplying each of these by 2. It appears that in each case, the result $2z$ is on the ray from the origin through z, and that $2z$ is twice as far from the origin as z is. We can see why this occurs by looking at the exponential or polar form of z. Writing z as $re^{i\theta}$, we have

$$2z = 2re^{i\theta}.$$

So, the magnitude of $2z$ is twice the magnitude of z, and the argument of $2z$ is the same as that of z. Because the arguments of $2z$ and z are the same, they are on the same ray from the origin. Therefore, we conclude that $2z$ is a dilation of z about the origin with scale factor 2.



(b) Our previous part gives us a clear guide. Writing z in exponential form gives us $cz = cre^{i\theta}$, so the magnitude of cz is c times the magnitude of z, and the argument of cz is the same as that of z. Therefore, cz is on the ray from the origin through z, and is c times as far from the origin as z, which means cz is a dilation of z about the origin with scale factor c.

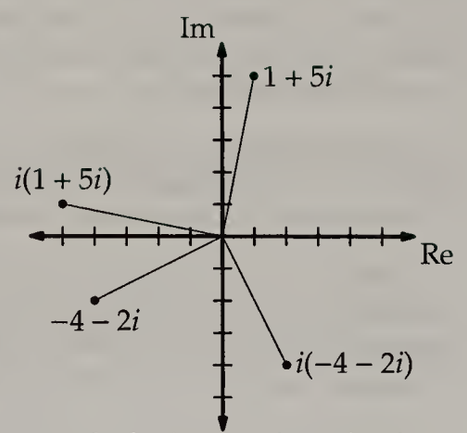
□

If c is negative, then the transformation that maps z to cz is still a dilation; it just has a negative scale factor, which means that cz is in the opposite direction from the origin from z.

Problem 8.5:
(a) How are z and iz related to each other in the complex plane?
(b) Generalize your result from part (a) by finding a relationship between z and $e^{i\alpha}z$ in the complex plane.

Solution for Problem 8.5:

(a) First, we experiment with a few values of z. If $z = 1$, then $iz = i$. If $z = 1+5i$, then $iz = i(1+5i) = -5+i$. If $z = -4-2i$, then $iz = i(-4-2i) = 2 - 4i$. The latter two examples are depicted at right. In each of these three cases, we see that iz is a $\frac{\pi}{2}$ counterclockwise rotation of z about the origin.



Now that we know what to look for, we have a pretty strong clue how to prove it. Angles and complex numbers suggest using exponential form. Letting $z = re^{i\theta}$ and noting that $i = e^{\pi i/2}$, we have

$$iz = e^{\frac{\pi i}{2}} \cdot re^{i\theta} = re^{i\left(\theta+\frac{\pi}{2}\right)}.$$

Therefore, iz is just as far from the origin in the complex plane as z is, but its argument is $\frac{\pi}{2}$ greater. That is, iz in the complex plane is a $\frac{\pi}{2}$ counterclockwise rotation of z about the origin.

(b) The previous part gives us a guide. We let $z = re^{i\theta}$, and we have

$$e^{i\alpha}z = e^{i\alpha} \cdot re^{i\theta} = re^{i(\alpha+\theta)}.$$

So, $e^{i\alpha}z$ and z are the same distance from the origin, but the argument of $e^{i\alpha}z$ is α greater than the argument of z. Therefore, $e^{i\alpha}z$ is the result of rotating z counterclockwise about the origin by an angle of α.

□

Important: When z is rotated α counterclockwise about the origin in the complex plane, the result is $e^{i\alpha}z$.

In Problems 8.1, 8.4, and 8.5, we explored a specific case of a transformation and then used our specific case as guidance to find a more general result. This is a very common and important strategy for finding proofs of general statements.

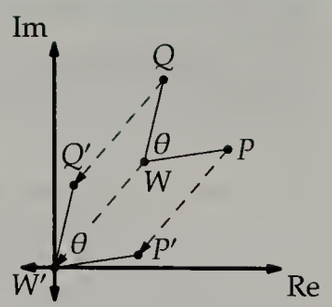
Concept: Many general results can be discovered by first exploring simple examples.

Problem 8.6: Describe in words the transformation that maps z to $re^{i\theta} \cdot z$, where r is a positive real constant and θ is a real constant.

Solution for Problem 8.6: The transformation that maps z to $re^{i\theta} \cdot z$ is simply a combination of the transformations we learned in Problem 8.4 and Problem 8.5. From Problem 8.5, we know that $e^{i\theta}z$ is the result of rotating z an angle of θ counterclockwise about the origin. From Problem 8.4, we know that $r \cdot (e^{i\theta}z)$ is the result of dilating $e^{i\theta}z$ by a scale factor of r from the origin. Therefore, the transformation that maps z to $re^{i\theta} \cdot z$ is a counterclockwise rotation by an angle θ followed by a dilation with scale factor r. □

Problem 8.7: If q is the result of rotating p an angle θ counterclockwise about w in the complex plane, then find q in terms of θ, p, and w.

Solution for Problem 8.7: Let Q be the image of a counterclockwise rotation of P about W by θ, as shown, and let points P, Q, and W correspond to complex numbers p, q, and w, respectively. The only rotations we know how to handle are rotations about the origin. So, we turn this problem into a problem we know how to handle by translating the center of rotation, W, to the origin. We apply the same translation to P and to Q, and we have the diagram at right, where P', Q', and W' are the images of P, Q, and W under the translation that maps W to the origin.



The translation that maps W to the origin must map the complex number w to 0. Therefore, this translation maps any complex number z to $z - w$. Since P' is the image of P under this translation, P' corresponds to the complex number $p' = p - w$. Similarly, we have $q' = q - w$, where q' corresponds to Q'.

Next, we note that since Q is the image of P upon rotation about W, we have $QW = WP$. Since translating a segment does not change its length, we have $Q'W' = W'P'$ as well. Translating an angle does not change its measure, so $\angle Q'W'P' = \angle QWP = \theta$. Since $Q'W' = W'P'$ and $\angle Q'W'P' = \theta$, point Q' is the image of point P' upon a counterclockwise rotation of angle θ.

The image of z upon a counterclockwise rotation by θ about the origin is $e^{i\theta}z$. Since Q' is the image of P' under such a rotation, we have $q' = e^{i\theta}p'$. Substituting our expressions above for p' and q', we have $q - w = e^{i\theta}(p - w)$.

Solving for q gives

$$q = e^{i\theta}(p - w) + w.$$

□

Important: If q is the result of rotating p about w in the complex plane counterclockwise by an angle of θ, then

$$q = e^{i\theta}(p - w) + w.$$

This is often easier to visualize in the form

$$q - w = e^{i\theta}(p - w),$$

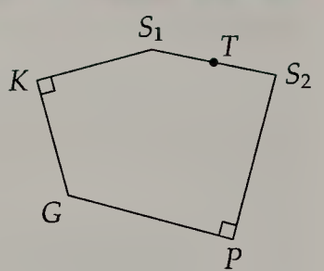
for the reasons we discussed in our solution to Problem 8.7: $q - w$ and $p - w$ are the images under the translation that maps w to the origin, and $e^{i\theta}(p - w)$ is the image upon rotating $p - w$ by θ about the origin.

Problem 8.8: A pirate buries his treasure on a deserted island with an oak tree, a pine tree, and a gallows. To choose the burial location, he starts at the gallows and walks to the oak tree, counting his steps. When he reaches the oak tree, he turns right 90 degrees, walks the same number of steps as from the gallows to the oak tree, and places a spike in the ground.

He returns to the gallows, and then walks to the pine tree, again counting his steps. When he reaches the pine tree, he turns left 90 degrees, walks the same number of steps as from the gallows to the pine tree, and places another spike in the ground. He then buries the treasure at the midpoint of the segment with the spikes as endpoints, and then removes the spikes from the ground.

Years later, the pirate returns to the island to get his treasure. Unfortunately, the gallows are gone, and there is no sign of where it had been. Fortunately, the oak tree and pine tree are still there. How can the pirate find his treasure?

Solution for Problem 8.8: We let G, P, and K be the gallows, the pine, and the oak, respectively. We also let S_1 and S_2 be the spikes, and T be the treasure, as shown in the diagram.



Because $GK = KS_1$ and $\angle GKS_1 = 90°$, the first spike is a 90° counterclockwise rotation of the gallows about the oak tree. (It's a counterclockwise rotation because the pirate turns right when he reaches the oak tree from the gallows.) We know how to handle rotations on the complex plane. Moreover, 90° rotations are particularly easy to handle with complex numbers. So, we place the problem on the complex plane. We let the lowercase letters of each of the point labels be the corresponding complex numbers.

Because S_1 is a 90° counterclockwise rotation of G about K, we have

$$s_1 = k + (g - k)e^{\pi i/2} = k + (g - k)i.$$

Similarly, the second spike is a 90° clockwise rotation (because the pirate turns *left* at the pine) of the gallows about the pine, so

$$s_2 = p + (g - p)e^{-\pi i/2} = p + (g - p)(-i).$$

To find the treasure, we have to figure out how to find the midpoint of the segment connecting two points on the complex plane. Fortunately, this is pretty easy. The midpoint of the segment connecting (x_1, y_1) and (x_2, y_2) in the Cartesian plane is $\left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right)$. In the complex plane, this means that the midpoint of the segment connecting $z_1 = x_1 + y_1 i$ and $z_2 = x_2 + y_2 i$ is

$$\left(\frac{x_1 + x_2}{2}\right) + \left(\frac{y_1 + y_2}{2}\right) i,$$

which equals $(z_1 + z_2)/2$. Intuitively, this makes sense—it tells us that the midpoint of the segment connecting two complex numbers in the complex plane is the average of the two numbers.

Applying this to our problem to find t, the complex number corresponding to the treasure, we have

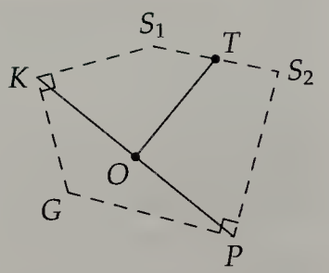
$$t = \frac{s_1 + s_2}{2} = \frac{k + (g-k)i + p + (g-p)(-i)}{2} = \frac{k + gi - ki + p - gi + pi}{2} = \frac{k+p}{2} + \frac{p-k}{2} \cdot i.$$

There's no g in this expression—that means that the location of the treasure doesn't depend on the gallows at all! It only depends on the pine and oak trees, which means that the pirate can still find the treasure.

But how? How can he find the point that corresponds to $\frac{k+p}{2} + \frac{p-k}{2} \cdot i$? One quick way he can do that is to choose his origin wisely.

Concept: When placing a geometric problem on the complex plane, we can choose where to place the origin. Choosing a convenient point can often simplify the problem greatly.

We recognize the first term in the expression $\frac{k+p}{2} + \frac{p-k}{2} \cdot i$; the point corresponding to $\frac{k+p}{2}$ is the midpoint of the segment connecting the trees. So, if we choose this midpoint to be the origin, then $\frac{k+p}{2} = 0$, and the treasure is simply at $\frac{p-k}{2} \cdot i$. Moreover, since $\frac{k+p}{2} = 0$, we have $k + p = 0$, so $k = -p$. Therefore, the treasure is at $\frac{p-(-p)}{2} \cdot i$, which simplifies to pi. Aha! The point corresponding to pi is a 90° counterclockwise rotation of p about the origin.



We now have a plan to find the treasure. We find the midpoint of the segment connecting the trees. We measure the distance from that point to the pine tree. We stand at the midpoint (the origin of our complex plane), face the pine tree, turn 90° counterclockwise, and then walk that measured distance in the direction we face. Then, we dig.

And we find the treasure. □

On the way to finding the treasure, we also discovered an unsurprising geometric relationship on the complex plane.

Important: The midpoint of the segment connecting z_1 and z_2 in the complex plane is $\frac{z_1+z_2}{2}$.

Exercises

8.1.1 Suppose w is the image of z upon a 180° rotation about the origin. Express w in terms of z.

8.1.2

(a) What is the image of $3 + i$ upon a 30° counterclockwise rotation about the origin?

(b) What is the image of $3 + i$ upon a 120° clockwise rotation about $1 + 2i$?

8.1.3 Note that argument of $w + z$ is average of the arguments of w and z in Problem 8.2. Is this always true? If not, when is it true?

8.1.4 Describe in words the transformation that maps z to $z/(re^{i\theta})$, where r is a positive real constant and θ is a real constant.

8.1.5 For each of the following functions, describe the corresponding transformation that maps z to $f(z)$.

(a) $f(z) = z - 5 + 2i$.

(b) $f(z) = 3z - 1$.

(c) $f(z) = \left(\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) z$.

(d) $f(z) = (-1 + i)z$.

(e) $f(z) = -iz - 1 + 3i$.

8.1.6 Suppose that U is on $\overline{WZ}$ in the complex plane such that $UW/UZ = a/b$. Show that if u, w, and z are complex numbers corresponding to U, W, and Z, then $u = (az + bw)/(a + b)$.

8.1.7⋆ Let z_1 be a complex number. Let z_2 be the complex number obtained when z_1 is rotated counterclockwise about the origin by $\frac{\pi}{2}$, and let z_3 be the complex number obtained when z_2 is rotated counterclockwise about $4 + 3i$ by $\frac{\pi}{2}$. Then z_3 is the same as the point obtained when z_1 is rotated counterclockwise about w by θ. Find w and θ.

8.2 Parallel and Perpendicular Lines

Problems

Problem 8.9: Let W and Z be points in the complex plane corresponding to nonzero complex numbers w and z, respectively, and let O be the origin.

(a) Show that if $\overleftrightarrow{WZ}$ passes through O, then z/w is a real number.

(b) Show that if z/w is a real number, then $\overleftrightarrow{WZ}$ must pass through O.

Problem 8.10: Suppose S, T, and U are points in the complex plane corresponding to distinct complex numbers s, t, and u, respectively. Three points are said to be **collinear** if a line passes through all three points. In this problem, we show that these three points are collinear if and only if $(s - t)/(u - t)$ is real.

Let O be the origin. Consider the translation that maps T to the origin. Let S' and U' be the images of S and U, respectively, under this translation, and let the complex numbers s' and u' correspond to S' and U', respectively.

(a) Explain why S, T, and U are collinear if and only if S', O, and U' are. What must be true of s' and u' if and only if S', O, and U' are collinear?

(b) Show that S, T, and U are collinear if and only if $(s - t)/(u - t)$ is real.

Problem 8.11: Let $\overline{AB}$ and $\overline{CD}$ be segments in the complex plane, and let a, b, c, and d be complex numbers corresponding to the endpoints of these segments. Show that $\overleftrightarrow{AB} \parallel \overleftrightarrow{CD}$ if and only if $(b - a)/(d - c)$ is real.

Problem 8.12: Use complex numbers to show that a quadrilateral is a parallelogram if and only if its diagonals bisect each other.

Problem 8.13: Let $\overline{AB}$ and $\overline{CD}$ be segments in the complex plane, and let a, b, c, and d be complex numbers corresponding to the endpoints of these segments. Find a condition similar to that in Problem 8.11 that is necessary and sufficient to show that $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$.

Problem 8.14: Let $\overline{AB}$ and $\overline{CD}$ be segments in the complex plane, and let a, b, c, and d be complex numbers corresponding to the endpoints of these segments.

(a) Show that $\overleftrightarrow{AB} \parallel \overleftrightarrow{CD}$ if and only if $(a-b)(\overline{c}-\overline{d}) - (\overline{a}-\overline{b})(c-d) = 0$.

(b) Show that $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$ if and only if $(a-b)(\overline{c}-\overline{d}) + (\overline{a}-\overline{b})(c-d) = 0$.

Problem 8.15: Suppose $VWXYZ$ is a pentagon such that the altitudes from V, W, X, and Y to the respective opposite sides of the pentagon (V to $\overline{XY}$, W to $\overline{YZ}$, etc.) meet at a common point P. In this problem, we show that the altitude from Z to $\overline{WX}$ passes through P as well.

(a) Write four equations corresponding to the four given altitudes. (Tip: Choose your origin wisely!)

(b) Write an equation that must be true if the altitude from Z to $\overline{WX}$ passes through P.

(c) Combine your equations in part (a) to produce the equation in part (b).

Problem 8.9: Let W and Z be points in the complex plane corresponding to nonzero complex numbers w and z, respectively, and let O be the origin. Show that $\overleftrightarrow{WZ}$ passes through O if and only if z/w is a real number.

Solution for Problem 8.9: We must show both of the following:

(a) If $\overleftrightarrow{WZ}$ passes through O, then z/w is real.

(b) If z/w is real, then $\overleftrightarrow{WZ}$ passes through O.

Important: Suppose $\mathcal{P}$ and $\mathcal{Q}$ are mathematical statements. The statement "$\mathcal{P}$ if and only if $\mathcal{Q}$" means both of the following:

- If $\mathcal{P}$ is true, then $\mathcal{Q}$ is also true.
- If $\mathcal{Q}$ is true, then $\mathcal{P}$ is also true.

Both of these statements must be proved in order to show that "$\mathcal{P}$ if and only if $\mathcal{Q}$."

We'll prove parts (a) and (b) separately. We'll start with part (a). Since W and Z are on the same line through the origin, we know that we can transform W to Z by performing a dilation about the origin (possibly with a negative scale factor). Therefore, there is some real constant c such that $z = cw$, from which we have $\frac{z}{w} = c$.

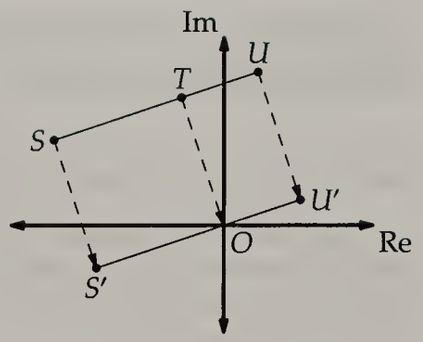
Next we must take care of part (b). If $\frac{z}{w} = c$, where c is a real number, then we have $z = cw$. Since c is real, this means that Z is the image of W under some dilation about the origin, which means that Z is on $\overleftrightarrow{OW}$, as desired.

We also could have tackled part (b) by noting that if $w = re^{i\theta}$, then $z = cre^{i\theta}$. If $c > 0$, then W and Z have the same argument. If $c < 0$, then their arguments differ by π. In both cases, W and Z lie on the same line through the origin. □

We say that three points are **collinear** if they lie on the same line. Now that we know how to tell if the line through two given points in the complex plane passes through the origin, let's see if we can develop a similar test to determine if any given three points are collinear.

Problem 8.10: Suppose S, T, and U are distinct points in the complex plane corresponding to complex numbers s, t, and u, respectively. Show that S, T, and U are collinear if and only if $(s-t)/(u-t)$ is real.

Solution for Problem 8.10: Just as we did with the general rotation problem in Problem 8.7, we tackle the general problem of collinearity by using translations to turn it into a problem we already know how to solve. We know how to tell if the line through two points passes through the origin, so we translate all three points such that the image of one of these points is the origin. The images of S, T, and U are collinear if and only if S, T, and U are. (We can see this by noting that applying a translation to $\angle STU$ will not change the measure of the angle.)



The expressions $s - t$ and $u - t$ in $(s - t)/(u - t)$ suggest translating T to the origin, since such a translation maps S and U to points S' and U' such that $s' = s - t$ and $u' = u - t$. Letting O be the origin, we can now apply our result from Problem 8.9. There, we showed that $\overleftrightarrow{S'U'}$ passes through the origin if and only if s'/u' is real. Substituting our expressions for s' and u', we see that $\overleftrightarrow{S'U'}$ passes through the origin if and only if $(s - t)/(u - t)$ is a real number. Since S', O, and U' are collinear if and only if $(s - t)/(u - t)$ is real, and S, T, and U are collinear if and only if S', O, and U' are collinear, we conclude that S, T, and U are collinear if and only if $(s - t)/(u - t)$ is real. □

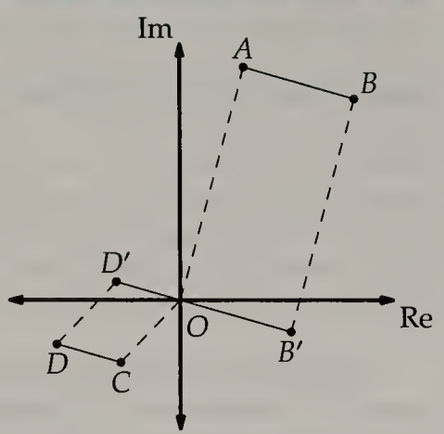
Important: The diagram at right above is not the only possible configuration; it is possible that T is on $\overleftrightarrow{SU}$, but not between S and U. However, the fact that $\overleftrightarrow{S'U'}$ passes through the origin if and only if $(s - t)/(u - t)$ is real *does not depend upon the possible orderings of S', U', and O.* So, we don't have to worry about different cases.

Important: Suppose S, T, and U are distinct points in the complex plane corresponding to complex numbers s, t, and u, respectively. S, T, and U are collinear if and only if $(s - t)/(u - t)$ is real.

We'll now turn to parallel lines. In addition to the usual definition that two lines are parallel if they do not intersect, we will also say that a line is parallel to itself. We can think of this as "all lines with the same direction are parallel." We will make the concept of "direction" of a line more rigorous in Section 9.3.

Problem 8.11: Let $\overline{AB}$ and $\overline{CD}$ be segments in the complex plane, and let a, b, c, and d be complex numbers corresponding to the endpoints of these segments. Show that $\overleftrightarrow{AB} \parallel \overleftrightarrow{CD}$ if and only if $(b - a)/(d - c)$ is a real number.

Solution for Problem 8.11: There's no need to reinvent the wheel here; we have a pretty clear guide how to tackle this problem in our solution to Problem 8.10. We start by translating $\overline{AB}$ such that the image of A is the origin, O. Let the image of B under this translation be B', so that $b' = b - a$. Similarly, we translate $\overline{CD}$ such that the image of C is the origin and the image of D is D', which gives us $d' = d - c$. Since $\overleftrightarrow{OB'} \parallel \overleftrightarrow{AB}$ and $\overleftrightarrow{OD'} \parallel \overleftrightarrow{CD}$, we have $\overleftrightarrow{AB} \parallel \overleftrightarrow{CD}$ if and only if $\overleftrightarrow{OB'}$ and $\overleftrightarrow{OD'}$ are the same line, because there is only one line through O parallel to $\overleftrightarrow{AB}$. In Problem 8.9, we saw that $\overleftrightarrow{OB'}$ and $\overleftrightarrow{OD'}$ are the same line if and only if b'/d' is real. Substituting our expressions above for b' and d', this tells us that $\overleftrightarrow{OB'}$ and $\overleftrightarrow{OD'}$ are the same line if and only if $(b - a)/(d - c)$ is real. Every step in our reasoning is an "if and only if" step, so we have proved that $\overleftrightarrow{AB} \parallel \overleftrightarrow{CD}$ if and only if $(b - a)/(d - c)$ is a real number. □



Important: Let $\overline{AB}$ and $\overline{CD}$ be segments in the complex plane, and let a, b, c, and d be complex numbers corresponding to the endpoints of these segments. Then lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ are parallel if and only if $(b - a)/(d - c)$ is a real number.

Problem 8.12: Use complex numbers to show that a quadrilateral is a parallelogram if and only if its diagonals bisect each other.

Solution for Problem 8.12: Let complex numbers a, b, c, and d correspond to the vertices of quadrilateral $ABCD$ in the complex plane. Then, the midpoint of diagonal $\overline{AC}$ is $\frac{a+c}{2}$ and the midpoint of diagonal $\overline{BD}$ is $\frac{b+d}{2}$.

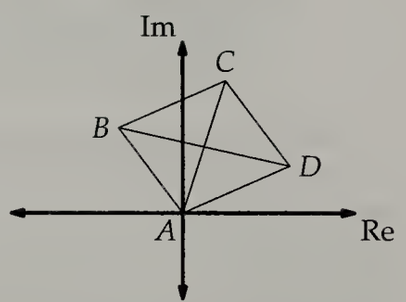
First, we show that if the diagonals of $ABCD$ bisect each other, then $ABCD$ is a parallelogram. If the diagonals of $ABCD$ bisect each other, then the midpoints of $\overline{AC}$ and $\overline{BD}$ are the same. That is, we have

$$\frac{a+c}{2} = \frac{b+d}{2},$$

so $a + c = b + d$. We'd like to show that $\overline{AB} \parallel \overline{DC}$, which we can prove by showing that $(b-a)/(c-d)$ is real. So, we rearrange $a + c = b + d$ as $c - d = b - a$. Dividing both sides by $c - d$ gives $\frac{b-a}{c-d} = 1$. Since this is real, we have $\overline{AB} \parallel \overline{DC}$. We could go through essentially the same steps to show that $\overline{AD} \parallel \overline{BC}$, or we could note that since $c - d = b - a$, we have $|c - d| = |b - a|$, which means that $DC = AB$. Combining $AB = DC$ with $\overline{AB} \parallel \overline{DC}$ tells us that $ABCD$ is a parallelogram.

Important: Let $\overline{AB}$ and $\overline{CD}$ be segments in the complex plane, and let a, b, c, and d be complex numbers corresponding to the endpoints of these segments. We have $AB = CD$ if and only if $|b - a| = |d - c|$.

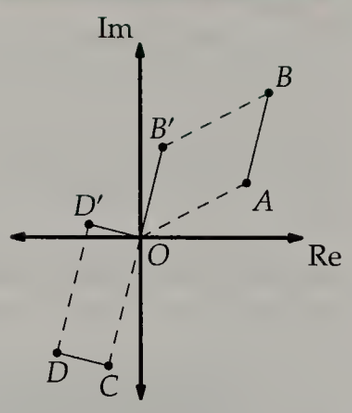
Next, we show that if $ABCD$ is a parallelogram, then the diagonals bisect each other. We can take a bit of a shortcut here by letting A be the origin, which means the midpoint of $\overline{AC}$ is simply $\frac{c}{2}$. Then, because $\overline{AB} \parallel \overline{DC}$ and $AB = DC$, the translation that maps B to A also maps C to D. This translation maps B to the origin, so it maps any complex number z to $z - b$. Because this translation also maps C to D, we have $d = c - b$. Therefore, we have $d + b = c$, so the midpoint of $\overline{BD}$ is $\frac{b+d}{2} = \frac{c}{2}$. This tells us that diagonals $\overline{AC}$ and $\overline{BD}$ have the same midpoint, which means they bisect each other. □



Having tackled parallel lines, let's turn to perpendicular lines.

Problem 8.13: Let $\overline{AB}$ and $\overline{CD}$ be segments in the complex plane, and let a, b, c, and d be complex numbers corresponding to the endpoints of these segments. Find a condition similar to that in Problem 8.11 that is necessary and sufficient to show that $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$.

Solution for Problem 8.13: Solution 1: Use the method of previous problems. As before, we start with translations to the origin. We translate $\overline{AB}$ such that the image of A is the origin. This translation maps B to a point B' that corresponds to $b' = b - a$. Similarly, we translate $\overline{CD}$ such that the image of C is the origin, which means the image of D is the point D' that corresponds to $d' = d - c$.



Translating a segment does not change its orientation, so $\overleftrightarrow{AB} \parallel \overleftrightarrow{OB'}$. Therefore, the angle $\overleftrightarrow{OD'}$ makes with $\overleftrightarrow{OB'}$ is the same as the angle $\overleftrightarrow{OD'}$ makes with $\overleftrightarrow{AB}$. Similarly, $\overleftrightarrow{OB'}$ makes the same angle with $\overleftrightarrow{CD}$ as it makes with $\overleftrightarrow{OD'}$. Combining these observations, we see that the angle between $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ is the same as that between $\overleftrightarrow{OB'}$ and $\overleftrightarrow{OD'}$. This means that $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$ if and only if $\overleftrightarrow{OB'} \perp \overleftrightarrow{OD'}$. Therefore, our problem now is to prove that $\overleftrightarrow{OB'} \perp \overleftrightarrow{OD'}$ if and only if $\frac{b-a}{d-c}$ is imaginary. Since $b' = b - a$ and $d' = d - c$, we must show that $\overleftrightarrow{OB'} \perp \overleftrightarrow{OD'}$ if and only if b'/d' is imaginary.

If $\overline{OB'} \perp \overline{OD'}$, then the arguments of b' and d' differ by $\frac{\pi}{2}$. Therefore, if d' in exponential form is $re^{i\theta}$, then b' is $se^{i(\theta+(\pi/2))}$ or $se^{i(\theta-(\pi/2))}$, where r and s are real numbers. If $b' = se^{i(\theta+(\pi/2))}$, then we have

$$\frac{b'}{d'} = \frac{se^{i(\theta+(\pi/2))}}{re^{i\theta}} = \frac{s}{r}e^{\pi i/2} = \frac{s}{r}i.$$

Similarly, if $b' = se^{i(\theta-(\pi/2))}$, then we find $\frac{b'}{d'} = -\frac{s}{r}i$. In either case, $\frac{s}{r}$ is real, so $\frac{b'}{d'}$ is an imaginary number.

We must also show that if $\frac{b'}{d'} = ki$ for some real constant k, then $\overline{OB'} \perp \overline{OD'}$. Multiplying both sides by d' gives $b' = (kd')i$. Therefore, B' is a 90° counterclockwise rotation about the origin of the point corresponding to kd'. The point corresponding to kd' is a dilation of D' about the origin, so it is on $\overleftrightarrow{OD'}$. Since B' is a 90° rotation about the origin of a point on $\overleftrightarrow{OD'}$, we have $\overline{OB'} \perp \overline{OD'}$, as desired.

Solution 2: Use the result of Problem 8.11. We turn our problem about perpendicular lines into one about parallel lines by rotating one of the lines 90° about the origin. Let $\overleftrightarrow{C'D'}$ be the image of $\overleftrightarrow{CD}$ upon a 90° rotation counterclockwise about the origin. Therefore, we have $c' = ci$ and $d' = di$. We have $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$ if and only if $\overleftrightarrow{AB} \parallel \overleftrightarrow{C'D'}$. From Problem 8.11, we have $\overleftrightarrow{AB} \parallel \overleftrightarrow{C'D'}$ if and only if $\frac{b-a}{d'-c'}$ is a real number. That is, $\overleftrightarrow{AB} \parallel \overleftrightarrow{C'D'}$ if and only if $\frac{b-a}{d'-c'} = t$ for some real constant t. Substituting our expressions for c' and d', this equation becomes

$$\frac{b-a}{di-ci} = t.$$

Multiplying both sides by i gives $\frac{b-a}{d-c} = ti$, which is imaginary. Therefore, we have $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$ if and only if $\frac{b-a}{d-c}$ is an imaginary number, as before. □

Important: Let $\overline{AB}$ and $\overline{CD}$ be segments in the complex plane, and let a, b, c, and d be complex numbers corresponding to the endpoints of these segments. Then $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$ if and only if $(b-a)/(d-c)$ is an imaginary number.

Problem 8.14: Let A, B, C, and D be four different points in the complex plane, and let a, b, c, and d be complex numbers corresponding to these points.

(a) Show that $\overleftrightarrow{AB} \parallel \overleftrightarrow{CD}$ if and only if $(b-a)(\overline{d}-\overline{c}) - (\overline{b}-\overline{a})(d-c) = 0$.

(b) Show that $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$ if and only if $(b-a)(\overline{d}-\overline{c}) + (\overline{b}-\overline{a})(d-c) = 0$.

Solution for Problem 8.14:

(a) In Problem 8.11, we showed that $\overleftrightarrow{AB} \parallel \overleftrightarrow{CD}$ if and only if $(b-a)/(d-c)$ is real. Back on page 204, we showed that a complex number z is real if and only if $\overline{z} = z$, which we can write as $z - \overline{z} = 0$. So, because $\overleftrightarrow{AB} \parallel \overleftrightarrow{CD}$ if and only if $(b-a)/(d-c)$ is real, we must have $z - \overline{z} = 0$ for $z = (b-a)/(d-c)$:

$$\frac{b-a}{d-c} - \overline{\left(\frac{b-a}{d-c}\right)} = 0. \tag{8.1}$$

Since we have

$$\overline{\left(\frac{b-a}{d-c}\right)} = \frac{\overline{b-a}}{\overline{d-c}} = \frac{\overline{b}-\overline{a}}{\overline{d}-\overline{c}},$$

we can write Equation (8.1) as

$$\frac{b-a}{d-c} - \frac{\overline{b}-\overline{a}}{\overline{d}-\overline{c}} = 0.$$

Multiplying both sides by $d-c$ and by $\overline{d}-\overline{c}$, we have the desired

$$(b-a)(\overline{d}-\overline{c}) - (\overline{b}-\overline{a})(d-c) = 0.$$

(b) We can follow essentially the same steps as in the previous part. In Problem 8.13 we showed that $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$ if and only if $(b-a)/(d-c)$ is imaginary. Back on page 204, we learned that a complex number z is imaginary if and only if $\overline{z} = -z$, which we can rearrange as $z + \overline{z} = 0$. So, letting $z = (b-a)/(d-c)$, we have

$$\frac{b-a}{d-c} + \overline{\left(\frac{b-a}{d-c}\right)} = 0,$$

which we can write as

$$\frac{b-a}{d-c} + \frac{\overline{b}-\overline{a}}{\overline{d}-\overline{c}} = 0.$$

Multiplying both sides by $(d-c)(\overline{d}-\overline{c})$ gives

$$(b-a)(\overline{d}-\overline{c}) + (\overline{b}-\overline{a})(d-c) = 0.$$

□

Our work in Problem 8.14 allows us to write the conditions "$(b-a)/(d-c)$ is a real number" and "$(b-a)/(d-c)$ is an imaginary number" as equations involving only a, b, c, and d. These equations are often much more convenient to work with than the qualitative description of $(b-a)/(d-c)$ being real or imaginary.

Important: Let A, B, C, and D be four different points in the complex plane, and let a, b, c, and d be complex numbers corresponding to these points.

- We have $\overleftrightarrow{AB} \parallel \overleftrightarrow{CD}$ if and only if $(b-a)(\overline{d}-\overline{c}) - (\overline{b}-\overline{a})(d-c) = 0$.
- We have $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$ if and only if $(b-a)(\overline{d}-\overline{c}) + (\overline{b}-\overline{a})(d-c) = 0$.

These aren't really new observations; they are equations that follow from the conditions we found earlier under which $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ are either parallel or perpendicular.

Looking back at our solution to Problem 8.14, we actually proved something a little bit *stronger* than we were required to. We needed A and B to be different to define line $\overleftrightarrow{AB}$, and needed C and D be different to define $\overleftrightarrow{CD}$. But nowhere in our proof do we use the fact that A and B differ from C and D! We could have had B and D be the same point, and every step in the proof holds. Therefore, by letting B and D be the same point, we have a condition for A, B, and C to be collinear:

Important: Let A, B, and C be three different points in the complex plane, and let a, b, and c be complex numbers corresponding to these points. A, B, and C are collinear if and only if

$$(b-a)(\overline{b}-\overline{c}) - (\overline{b}-\overline{a})(b-c) = 0.$$

The condition for perpendicularity given above also holds if we let B and D be the same point.

Concept: Sometimes we can prove a statement by proving a more general statement.

Problem 8.15: Suppose $VWXYZ$ is a pentagon such that the altitudes from V, W, X, and Y to the respective opposite sides of the pentagon (V to $\overline{XY}$, W to $\overline{YZ}$, etc.) meet at a common point P. Show that the altitude from Z to $\overline{WX}$ passes through P as well.

Solution for Problem 8.15: We are given four pairs of perpendicular lines and must use them to prove a fifth pair of lines are perpendicular. One point is common to all the pairs of lines: P. So, we'll let P be the origin, expecting that to make our algebra simpler. We let v, w, x, y, z be the complex numbers corresponding to the respective vertices of the pentagon. Since $\overleftrightarrow{VP} \perp \overleftrightarrow{XY}$, we have

$$(v-0)(\overline{x}-\overline{y})+(\overline{v}-0)(x-y)=0.$$

We simplify $v-0$ and $\overline{v}-0$, and then write similar equations for each of the other three altitudes, and we have

$$\begin{aligned} v(\overline{x}-\overline{y})+\overline{v}(x-y)&=0,\\ w(\overline{y}-\overline{z})+\overline{w}(y-z)&=0,\\ x(\overline{z}-\overline{v})+\overline{x}(z-v)&=0,\\ y(\overline{v}-\overline{w})+\overline{y}(v-w)&=0. \end{aligned}$$

We'd like to prove that $\overleftrightarrow{ZP} \perp \overleftrightarrow{WX}$, which is true if and only if

$$z(\overline{w}-\overline{x})+\overline{z}(w-x)=0. \tag{8.2}$$

Looking at our initial four equations above, we see that all the terms in the expansion of Equation (8.2) appear in the expansions of the first four equations. Unfortunately, there are a bunch more terms in those four equations. Let's go ahead and expand them and get a closer look at those extra terms:

$$\begin{aligned} v\overline{x}-v\overline{y}+\overline{v}x-\overline{v}y&=0,\\ w\overline{y}-w\overline{z}+\overline{w}y-\overline{w}z&=0,\\ x\overline{z}-x\overline{v}+\overline{x}z-\overline{x}v&=0,\\ y\overline{v}-y\overline{w}+\overline{y}v-\overline{y}w&=0. \end{aligned}$$

If we add all of these equations, the extra terms will all cancel! For example, the $v\overline{x}$ in the first equation cancels with the $-\overline{x}v$ in the third equation, and the $-v\overline{y}$ in the first equation cancels with the $\overline{y}v$ in the fourth equation. The terms that don't cancel when we add all four equations leave us with

$$x\overline{z}-w\overline{z}+\overline{x}z-\overline{w}z=0.$$

Factoring, multiplying by -1, and doing a little rearranging, gives us

$$z(\overline{w}-\overline{x})+\overline{z}(w-x)=0,$$

so $\overleftrightarrow{ZP} \perp \overleftrightarrow{WX}$, as desired. □

We glossed over an important strategy in setting up our solution to Problem 8.15. We wrote the four equations corresponding to the four given altitudes in a way that made it clear that we would have a lot of cancellation when we added the four equations. For example, if instead of writing the four equations

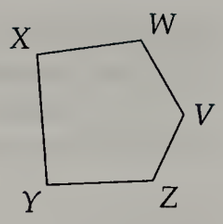
$$\begin{aligned} v(\overline{x}-\overline{y})+\overline{v}(x-y)&=0,\\ w(\overline{y}-\overline{z})+\overline{w}(y-z)&=0,\\ x(\overline{z}-\overline{v})+\overline{x}(z-v)&=0,\\ y(\overline{v}-\overline{w})+\overline{y}(v-w)&=0, \end{aligned}$$

we had written the four equations

$$\begin{aligned} v(\overline{x}-\overline{y})+\overline{v}(x-y)&=0,\\ w(\overline{y}-\overline{z})+\overline{w}(y-z)&=0,\\ x(\overline{z}-\overline{v})+\overline{x}(z-v)&=0,\\ y(\overline{w}-\overline{v})+\overline{y}(w-v)&=0, \end{aligned}$$

then adding the four equations won't cancel the $v\overline{y}$ and $y\overline{v}$ terms. We'd have to subtract the last equation instead, and our key step would be "add the first three equations and subtract the last," which is a much harder key step to see. (Look at the last equation in each set to see the difference between the two systems.)

We make sure our equations are set up in a convenient way by building them in a consistent manner. Here, we can set our equations up with a nice symmetry by visualizing the pentagon, and making sure the three points in each equation appear in the same order that they appear when we travel counterclockwise around the pentagon. For example, we go counterclockwise when we go from V to X to Y around the pentagon, and v, x, and y appear in that order in the first equation. Similarly, we go counterclockwise to go from W to Y to Z, so w, y, z appear in that order in the second equation, and so on.



Concept: Maintaining symmetry in systems of equations helps us find effective ways to work with the equations.

Exercises

8.2.1 Which of the following complex numbers lies on the line through $1 + 3i$ and $5 - 2i$ in the complex plane: $3 - 2i$, $9 - 7i$, $-2 + 5i$.

8.2.2 Show that the equation whose graph consists of all complex numbers z on the line through a and b in the complex plane is

$$\frac{z-a}{b-a} = \frac{\bar{z}-\bar{a}}{\bar{b}-\bar{a}}.$$

8.2.3 Prove that the quadrilateral formed by connecting the midpoints of any quadrilateral is a parallelogram.

8.2.4 In Problem 8.13, we used an algebraic approach to show that if O is the origin and points B' and D' in the complex plane satisfy $\overline{OB'} \perp \overline{OD'}$, then $\frac{b'}{d'}$ is an imaginary number, where b' and d' are the complex numbers corresponding to B' and D', respectively. Prove this fact using geometric transformations. That is, find a sequence of transformations that maps B' to D', and use this sequence to explain why $\frac{b'}{d'}$ must be imaginary.

8.2.5 A **median** of a triangle is the segment from a vertex of the triangle to the midpoint of the opposite side. In this problem, we show that the medians of a triangle are concurrent at a point called the **centroid**, and show that the distance from a vertex of the triangle to the triangle's centroid is 2/3 the length of the median from that vertex.

(a) Let a, b, and c be complex numbers corresponding to the respective vertices of $\triangle ABC$. Show that if $g = (a + b + c)/3$ corresponds to point G, then point G is on all three medians.

(b) Let M be the midpoint of $\overline{BC}$. Show that $AG/GM = 2$.

8.2.6 Show that the complex numbers a, b, and c are collinear if and only if

$$a\bar{b} + b\bar{c} + c\bar{a} = \bar{a}b + \bar{b}c + \bar{c}a.$$

8.2.7 In convex quadrilateral $WXYZ$, the midpoints of $\overline{WX}$, $\overline{WY}$, and $\overline{XZ}$ are collinear. Show that the line through these points also passes through the midpoint of $\overline{YZ}$.

8.2.8$\star$ Let A correspond to $1 - 5i$, B correspond to $3 + 2i$, C correspond to $-7\sqrt{3} + 7i$, and D correspond to $-2 - 2i\sqrt{3}$. Find the acute angle between $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$.

8.3 Distance

We introduced the magnitude of a complex number as the distance between that complex number and the origin in the complex plane. From this definition, we also saw that if w and z are complex numbers, then $|w - z|$ is the

distance between w and z in the complex plane. In this section, we use the magnitude of complex numbers to solve problems involving lengths of segments.

Problems

Problem 8.16: Let complex numbers w and z correspond to points W and Z in the complex plane. Show that $WZ^2 = (w-z)(\overline{w}-\overline{z})$.

Problem 8.17: Prove that $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$ if and only if $AC^2 + BD^2 = AD^2 + BC^2$.

Problem 8.18: Let c be a complex number and t be a real number such that $t < c\overline{c}$. In this problem, we show that the graph of an equation of the form

$$z\overline{z} - c\overline{z} - \overline{c}z + t = 0$$

is a circle.

(a) Find an equation whose graph is a circle with center c and radius r.

(b) Show that the equation you found in part (a) is equivalent to $z\overline{z} - c\overline{z} - \overline{c}z + t = 0$ for some real number t. Why must we have the restriction $t < c\overline{c}$?

Problem 8.19: In this problem, we find the figure that consists of all points Z such that $AZ = k \cdot BZ$, where A and B are points and k is a positive real number with $k \neq 1$.

(a) Let A be the origin and B be 1 in the complex plane. Find an equation whose graph consists of the points Z such that $AZ = k \cdot BZ$.

(b) Show that the desired figure is a circle.

Problem 8.20: The **Triangle Inequality** is a fancy way to say that the shortest distance between two points is a straight line. Specifically, the Triangle Inequality tells us that for any three points A, B, and C, we have

$$AB + BC \geq AC,$$

where equality holds if and only if B is on $\overline{AC}$.

(a) Let w and z be complex numbers. Use the Triangle Inequality to prove that $|w| + |z| \geq |w+z|$. Under what conditions do we have $|w| + |z| = |w+z|$?

(b) Let $z_1, z_2, \ldots, z_n$ be complex numbers. Prove that $|z_1| + |z_2| + \cdots + |z_n| \geq |z_1 + z_2 + \cdots + z_n|$. Under what conditions do we have $|z_1| + |z_2| + \cdots + |z_n| = |z_1 + z_2 + \cdots + z_n|$?

(c) Show that for any real numbers a, b, c, and d, we have $\sqrt{a^2+b^2} + \sqrt{c^2+d^2} \geq \sqrt{(a+c)^2 + (b+d)^2}$.

Problem 8.21: Find the smallest possible value of $|z-5| + |z+4-3i|$.

Problem 8.16: Let complex numbers w and z correspond to points W and Z in the complex plane. Show that $WZ^2 = (w-z)(\overline{w}-\overline{z})$.

Solution for Problem 8.16: We have $WZ = |w-z|$, so

$$WZ^2 = |w-z|^2 = (w-z)\overline{(w-z)} = (w-z)(\overline{w}-\overline{z}),$$

as desired. □

Important: If complex numbers w and z correspond to points W and Z in the complex plane, then $WZ^2 = (w-z)(\overline{w}-\overline{z})$.

Problem 8.17: Prove that $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$ if and only if $AC^2 + BD^2 = AD^2 + BC^2$.

Solution for Problem 8.17: We think to use complex numbers in this problem because we have a straightforward expression for the square of the distance between two points in the complex plane, and we have a somewhat similar simple condition for perpendicularity. We let a, b, c, and d be complex numbers corresponding to points A, B, C, and D in the complex plane. We then have

$$AC^2 + BD^2 = (a-c)(\overline{a}-\overline{c}) + (b-d)(\overline{b}-\overline{d}) = a\overline{a} - a\overline{c} - c\overline{a} + c\overline{c} + b\overline{b} - b\overline{d} - d\overline{b} + d\overline{d},$$
$$AD^2 + BC^2 = (a-d)(\overline{a}-\overline{d}) + (b-c)(\overline{b}-\overline{c}) = a\overline{a} - a\overline{d} - d\overline{a} + d\overline{d} + b\overline{b} - b\overline{c} - c\overline{b} + c\overline{c}.$$

We see a lot of common terms in these, so subtracting will give us some cancellation:

$$\begin{aligned}(AC^2 + BD^2) - (AD^2 + BC^2) &= (-a\overline{c} - c\overline{a} - b\overline{d} - d\overline{b}) + (a\overline{d} + d\overline{a} + b\overline{c} + c\overline{b}) \\ &= a(\overline{d}-\overline{c}) + b(\overline{c}-\overline{d}) + c(\overline{b}-\overline{a}) + d(\overline{a}-\overline{b}) \\ &= a(\overline{d}-\overline{c}) - b(\overline{d}-\overline{c}) - c(\overline{a}-\overline{b}) + d(\overline{a}-\overline{b}) \\ &= (a-b)(\overline{d}-\overline{c}) + (d-c)(\overline{a}-\overline{b}).\end{aligned}$$

So, we now have $AC^2 + BD^2 = AD^2 + BC^2$ if and only if $(a-b)(\overline{d}-\overline{c}) + (d-c)(\overline{a}-\overline{b}) = 0$. But this is exactly the condition for perpendicularity that we proved on page 270. Since $AC^2 + BD^2 = AD^2 + BC^2$ if and only if $(a-b)(\overline{d}-\overline{c}) + (d-c)(\overline{a}-\overline{b}) = 0$, and $(a-b)(\overline{d}-\overline{c}) + (d-c)(\overline{a}-\overline{b}) = 0$ if and only if $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$, we have $AC^2 + BD^2 = AD^2 + BC^2$ if and only if $AC^2 + BD^2 = AD^2 + BC^2$. □

So far, we've discussed points, lines, and distance. Since a circle is defined in terms of a point and a distance, we can describe a circle in the complex plane as the graph of an equation.

Problem 8.18: Let c be a complex number and t be a real number such that $t < c\overline{c}$. Show that the graph of an equation of the form

$$z\overline{z} - c\overline{z} - \overline{c}z + t = 0$$

is a circle.

Solution for Problem 8.18: If point z is on the circle with center c and radius r, then $|z-c| = r$. Since both sides of $|z-c| = r$ are positive, the graph of $|z-c| = r$ is the same as the graph of $|z-c|^2 = r^2$. We prefer this latter form because it allows us to get rid of the magnitude symbol, since $|z-c|^2 = (z-c)\overline{(z-c)} = (z-c)(\overline{z}-\overline{c})$. So, the equation $|z-c|^2 = r^2$ becomes

$$(z-c)(\overline{z}-\overline{c}) = r^2.$$

Expanding the left side and subtracting r^2 from both sides gives

$$z\overline{z} - \overline{c}z - c\overline{z} + c\overline{c} - r^2 = 0.$$

Since $c\overline{c}$ must be real and r^2 is real, letting $t = c\overline{c} - r^2$ gives us

$$z\overline{z} - \overline{c}z - c\overline{z} + t = 0,$$

where t is the real number $c\overline{c} - r^2$.

To see that the graph of any such equation is a circle if $t < c\bar{c}$, we need only walk backwards through our steps above. We write t as $c\bar{c} + t - c\bar{c}$, so we have

$$z\bar{z} - \bar{c}z - c\bar{z} + c\bar{c} + t - c\bar{c} = 0.$$

Factoring the first four terms and moving the other terms to the other side of the equation gives $(z-c)(\bar{z}-\bar{c}) = c\bar{c}-t$, so $(z-c)\overline{(z-c)} = c\bar{c} - t$. Therefore, we have

$$|z-c|^2 = c\bar{c} - t.$$

The left side must be nonnegative. Now we see why we must have the condition $t < c\bar{c}$. We can write $c\bar{c} - t = r^2$ for some positive r, and we have $|z-c|^2 = r^2$. Taking the square root gives $|z-c| = r$, and we know the graph of this equation is a circle with center c and radius r. □

The equation we just found for a circle in the complex plane is rather unwieldy, so we usually stick to the tools of Euclidean geometry in problems involving circles. However, applications of complex numbers to circles can be very effective, as we'll see in the next problem.

Problem 8.19: What figure consists of all points Z such that $AZ = k \cdot BZ$, where A and B are points and k is a positive real number with $k \neq 1$?

Solution for Problem 8.19: We let a, b, and z correspond to A, B, and Z, respectively, so that $AZ = k \cdot BZ$ becomes $|z-a| = k|z-b|$. Before we square this equation to get rid of the magnitudes (which we can safely do because both sides are positive), we make life a bit easier on ourselves by letting $a = 0$ and $b = 1$, because this makes our equation $|z| = k|z-1|$. We can make this simplification because no matter where A and B are, we can draw the axes such that A is the origin and B corresponds to 1 on the real axis.

Squaring both sides of $|z| = k|z-1|$ gives $|z|^2 = k^2|z-1|^2$, from which we have

$$z\bar{z} = k^2(z-1)\overline{(z-1)} = k^2(z-1)(\bar{z}-1) = k^2(z\bar{z} - z - \bar{z} + 1).$$

Rearranging gives

$$(k^2-1)z\bar{z} - k^2z - k^2\bar{z} + k^2 = 0.$$

Dividing by $k^2 - 1$ gives

$$z\bar{z} - \frac{k^2}{k^2-1}z - \frac{k^2}{k^2-1}\bar{z} + \frac{k^2}{k^2-1} = 0.$$

If we let $c = t = \frac{k^2}{k^2-1}$, then we can write this equation in the form that we found in Problem 8.18:

$$z\bar{z} - \bar{c}z - c\bar{z} + t = 0.$$

(Note that $c = \bar{c}$ because c is real.) We must only confirm that $c\bar{c} - t > 0$ to deduce that the graph is a circle. Since $t = c = \bar{c}$, we have

$$c\bar{c} - t = c^2 - c = c(c-1) = \left(\frac{k^2}{k^2-1}\right)\left(\frac{k^2}{k^2-1} - 1\right) = \left(\frac{k^2}{k^2-1}\right)\left(\frac{1}{k^2-1}\right) = \frac{k^2}{(k^2-1)^2},$$

which is clearly positive, so the graph is a circle. Since $c = \frac{k^2}{k^2-1}$, the center of this circle is the point corresponding to $\frac{k^2}{k^2-1}$ on the real axis. And since $c\bar{c} - t = \frac{k^2}{(k^2-1)^2}$, the radius of the circle is $\frac{k}{k^2-1}$. □

The circle we discovered in Problem 8.19 is called a **circle of Apollonius**, which is a circle that is defined as the set of points Z such that $AZ = k \cdot BZ$ for some given points A and B and some positive constant $k \neq 1$.

We'll finish our discussion of distance in the complex plane with the **Triangle Inequality**, which is a fancy way to say that the shortest distance between two points is a straight line.

Important: The Triangle Inequality states that for any three points A, B, and C, we have

$$AB + BC \geq AC,$$

where equality holds if and only if B is on $\overline{AC}$.

Problem 8.20:

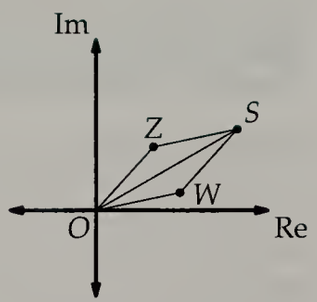
(a) Let w and z be complex numbers. Show that $|w| + |z| \geq |w + z|$. Under what conditions do we have $|w| + |z| = |w + z|$?

(b) Let $z_1, z_2, \ldots, z_n$ be complex numbers. Prove that $|z_1| + |z_2| + \cdots + |z_n| \geq |z_1 + z_2 + \cdots + z_n|$. Under what conditions do we have $|z_1| + |z_2| + \cdots + |z_n| = |z_1 + z_2 + \cdots + z_n|$?

(c) Show that for any real numbers a, b, c, and d, we have $\sqrt{a^2 + b^2} + \sqrt{c^2 + d^2} \geq \sqrt{(a + c)^2 + (b + d)^2}$.

Solution for Problem 8.20:

(a) First, we note that the inequality obviously holds (and equality holds) if w or z is 0. To address the case in which both are nonzero, we turn to the complex plane. Let points W, Z, and S correspond to complex numbers w, z, and $w + z$, respectively, and let O be the origin. As we have seen several times before, $OWSZ$ is a parallelogram, so $WS = OZ = |z|$. Applying the Triangle Inequality to $\triangle OWS$ gives us $OW + WS \geq OS$. We have $OW = |w|$, $WS = |z|$ and $OS = |w + z|$, so we have $|w| + |z| \geq |w + z|$, as desired.



We have equality in the Triangle Inequality $OW + WS \geq OS$ if and only if W is on $\overline{OS}$, so we have equality if and only if w is on the segment connecting the origin to $w + z$ in the complex plane. Since w is nonzero, this means $w + z$ must be nonzero and that $w = c(w + z)$ for some real constant c with $0 < c \leq 1$. Solving for z, we have $z = \frac{w}{c} - w = \left(\frac{1}{c} - 1\right) w$. Since $0 < c \leq 1$, we have $\frac{1}{c} - 1 > 0$, so $z = \left(\frac{1}{c} - 1\right) w$ tells us that z is a dilation of w about the origin with nonnegative scale factor. This means that w and c are on the same ray from the origin.

(b) We can extend the first part to the sum of the magnitudes of any number of complex numbers. Repeatedly applying part (a), we have

$$\begin{aligned} |z_1| + |z_2| + |z_3| + \cdots + |z_n| &\geq |z_1 + z_2| + |z_3| + \cdots + |z_n| && (8.3) \\ &\geq |z_1 + z_2 + z_3| + |z_4| + \cdots + |z_n| && (8.4) \\ &\geq \quad \vdots \\ &\geq |z_1 + z_2 + z_3 + \cdots + z_{n-1}| + |z_n| \\ &\geq |z_1 + z_2 + \cdots + z_n|. \end{aligned}$$

To see when equality holds, we note that we must have equality in each step above in order to have

$$|z_1| + |z_2| + |z_3| + \cdots + |z_n| = |z_1 + z_2 + \cdots + z_n|.$$

In our first step on line (8.3), we use the result of part (a) to get $|z_1| + |z_2| \geq |z_1 + z_2|$, where equality holds if and only if z_1 and z_2 are on the same ray through the origin. If z_1 and z_2 are on the same ray through the origin, then $z_1 + z_2$ is on this ray as well. In our next step on line (8.4), we again apply the result of part (a) to get $|z_1 + z_2| + |z_3| \geq |z_1 + z_2 + z_3|$, where equality holds if and only if $z_1 + z_2$ and z_3 are on the same ray through the origin. The ray from the origin through $z_1 + z_2$ is also the ray through z_1 and z_2 (if equality holds in our first application of the Triangle Inequality), so we have $|z_1| + |z_2| + |z_3| \geq |z_1 + z_2 + z_3|$ if and only if z_1, z_2, and z_3 are on the same ray from the origin. Continuing in this way, we see that we have

$$|z_1| + |z_2| + |z_3| + \cdots + |z_n| = |z_1 + z_2 + \cdots + z_n|$$

if and only if $z_1, z_2, z_3, \ldots, z_n$ are on the same ray from the origin.

(c) We can use magnitude to relate the square root of a sum of squares to a complex number: $|a+bi| = \sqrt{a^2+b^2}$. Similarly, we have $|c+di| = \sqrt{c^2+d^2}$, so from the Triangle Inequality expression

$$|a+bi| + |c+di| \geq |(a+bi)+(c+di)|,$$

we have

$$\sqrt{a^2+b^2} + \sqrt{c^2+d^2} \geq |(a+c)+(b+d)i| = \sqrt{(a+c)^2+(b+d)^2}.$$

□

Important: For any complex numbers $z_1, z_2, z_3, \ldots, z_n$, we have

$$|z_1| + |z_2| + \cdots + |z_n| \geq |z_1 + z_2 + \cdots + z_n|,$$

where equality holds if and only if $z_1, z_2, z_3, \ldots, z_n$ are on the same ray from the origin.

The version with two terms, $|z_1| + |z_2| \geq |z_1 + z_2|$, is often referred to as the **Triangle Inequality**. In Problem 8.20, we proved this algebraic expression of the Triangle Inequality by using the geometric expression of it. You might wonder if it's possible to go the other way—can we use the algebra of complex numbers to deduce that $|z_1| + |z_2| \geq |z_1 + z_2|$, and then use this to produce the geometric interpretation of the Triangle Inequality? Reversing our solution to part (a) above takes us from $|z_1| + |z_2| \geq |z_1 + z_2|$ to the geometric interpretation of the Triangle Inequality, but proving that $|z_1| + |z_2| \geq |z_1 + z_2|$ with only algebraic tools is a bit trickier! You'll have a chance to show how to do it as an Exercise.

Problem 8.21: Find the smallest possible value of $|z-5| + |z+4-3i|$.

Solution for Problem 8.21: Solution 1: Algebra. Applying the Triangle Inequality gives

$$|z-5| + |z+4-3i| \geq |(z-5)+(z+4-3i)| = |2z-1-3i|.$$

This isn't much help, since we can't do much with $|2z-1-3i|$.

We'd like to use the Triangle Inequality in a way that eliminates z. We could do so if the z term were negative in one of terms. We can make that happen! Since $|w| = |-w|$ for any w, we have

$$|z-5| + |z+4-3i| = |z-5| + |-(z+4-3i)| = |z-5| + |-z-4+3i| \geq |(z-5)+(-z-4+3i)| = |-9+3i| = 3\sqrt{10}.$$

We still have to confirm that this value is possible. We do so by noting that when $z = 5$, we have $|z-5| + |z+4-3i| = 0 + |9-3i| = 3\sqrt{10}$.

Important: Finding the minimum value of an expression is a two-step process. We must show both that the minimum value is achievable, and that no smaller value is possible. (Similarly, finding the maximum value of an expression is also a two-step process.)

The key step here was thinking about what we'd like to have happen:

Concept: A little wishful thinking can go a long way.

Solution 2: Geometry. The expression $|z-5|$ is the distance from z to 5 in the complex plane. The expression $|z+4-3i|$ equals $z-(-4+3i)$, which is the distance from z to $-4+3i$ in the complex plane. By the Triangle Inequality, the sum of the distances from z to 5 and to $-4+3i$ in the complex plane is at least the distance from 5 to $-4+3i$. The distance between these two points is $|5-(-4+3i)| = |9-3i| = 3\sqrt{10}$. As before, we must show that this value is achievable. From the Triangle Inequality, we know that we achieve this value if z is any point on the segment from 5 to $-4+3i$. We can obviously take z to be either endpoint, and we'll get $3\sqrt{10}$ as the value of the given expression. □

Our second solution reinforces an important strategy when working with complex numbers:

Concept: Some complex number problems that don't appear to be about geometry can be solved quickly by thinking about the problem geometrically.

Exercises

8.3.1 What is the radius of the circle that is the graph of the equation $z\bar{z} + (-2+i)z + (-2-i)\bar{z} = 11$?

8.3.2 Show that if P is a point on altitude $\overline{AX}$ of $\triangle ABC$, then $AC^2 - CP^2 = AB^2 - BP^2$.

8.3.3

(a) Prove that $|x+y-z| + |x-y+z| \geq 2|x|$.

(b) Prove that $|x| + |y| + |z| \leq |x+y-z| + |x-y+z| + |-x+y+z|$.

8.3.4 Show that if w is a complex number and a and b are nonzero real numbers such that $|w| = |w+a| = |w+bi| = 1$, then $a^2 + b^2 = 4$.

8.3.5 Show that the sum of the lengths of the diagonals of a quadrilateral is less than the perimeter of the quadrilateral.

8.3.6⋆ In the text, we used the geometric version of the Triangle Inequality to prove a version of the Triangle Inequality for complex numbers, $|w| + |z| \geq |w+z|$. Prove this inequality using only the algebraic properties of complex numbers, without using geometry. **Hints:** 93, 288, 87

8.4 Regular Polygons

We've already seen that complex numbers can be related to regular polygons when we discovered that plotting the n^{th} roots of unity on the complex plane gives us the vertices of a regular n-gon. In this section, we explore geometry problems involving regular polygons that can be tackled with complex numbers.

Problems

Problem 8.22: Suppose w and z are the endpoints of a diagonal of a square in the complex plane. Find the other two vertices of the square in terms of w and z.

Problem 8.23: The points $(0,0)$, $(a,11)$, and $(b,37)$ are the vertices of an equilateral triangle. Find ab. *(Source: AIME)* **Hints:** 80

Problem 8.24: Let ABC be a triangle in the complex plane whose vertices correspond to the complex numbers a, b, and c. Prove that $\triangle ABC$ is equilateral if and only if $a-b = \zeta(c-b)$, where ζ is a primitive sixth root of unity.

Problem 8.25: Suppose $\triangle P_1P_2P_3$ and $\triangle Q_1Q_2Q_3$ are equilateral triangles with the same orientation. (By same orientation, we mean that the vertices are labeled in the same order, either clockwise or counterclockwise, in each triangle.) Let M_1, M_2, and M_3 be the midpoints of $\overline{P_1Q_1}$, $\overline{P_2Q_2}$, and $\overline{P_3Q_3}$, respectively. Show that $\triangle M_1M_2M_3$ is an equilateral triangle.

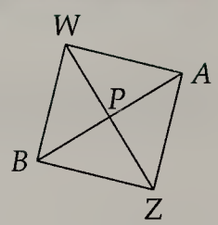
Problem 8.26: Points X and Y are outside $\triangle ABC$ such that $\triangle XAB$ and $\triangle YBC$ are equilateral and do not overlap $\triangle ABC$. Let P be the midpoint of $\overline{XB}$, Q be the midpoint of $\overline{YB}$, and R be the midpoint of $\overline{AC}$. Prove that $\triangle PQR$ is equilateral.

Problem 8.27: Let $n \geq 3$ be a positive integer, and let $P_0P_1 \cdots P_{n-1}$ be a regular n-gon inscribed in a circle with radius 1. Compute $P_0P_1 \cdot P_0P_2 \cdots P_0P_{n-1}$.

We'll start with a warm-up by considering a regular polygon that's particularly easy to work with: a square.

Problem 8.22: Suppose w and z are the endpoints of a diagonal of a square in the complex plane. Find the other two vertices of the square in terms of w and z.

Solution for Problem 8.22: Let complex numbers w and z correspond to points W and Z in the complex plane, and let A and B be the other vertices of the square, as shown at right. While it's not immediately obvious how to find A and B from W and Z, if we could locate the center of the square, we could quickly find A and B by rotating W and Z by 90° about this center. But we know how to find the center—it's the midpoint of diagonal $\overline{WZ}$.



Letting this midpoint be P, with corresponding complex number p, we have $p = \frac{w+z}{2}$. Since $AP = WP = BP = ZP$ and $\overline{AB} \perp \overline{WZ}$, points A and B are 90° rotations of W and Z about P. The complex number corresponding to a 90° counterclockwise rotation of Z about P is

$$p + (z-p)e^{\pi i/2} = \frac{w+z}{2} + \left(z - \frac{w+z}{2}\right) i = \left(\frac{1}{2} - \frac{i}{2}\right) w + \left(\frac{1}{2} + \frac{i}{2}\right) z,$$

and the complex number corresponding to a 90° counterclockwise rotation of W about P is

$$p + (w-p)e^{\pi i/2} = \frac{w+z}{2} + \left(w - \frac{w+z}{2}\right) i = \left(\frac{1}{2} + \frac{i}{2}\right) w + \left(\frac{1}{2} - \frac{i}{2}\right) z.$$

We can let either of these correspond to A, and the other corresponds to B. □

Problem 8.23: The points $(0,0)$, $(a,11)$, and $(b,37)$ are the vertices of an equilateral triangle. Find ab. *(Source: AIME)*

Solution for Problem 8.23: Let O be the origin, A be $(a,11)$, and B be $(b,37)$. We could start out with the distance formula. Since $OA = AB = BO$, we have

$$\sqrt{a^2 + 11^2} = \sqrt{(a-b)^2 + (11-37)^2} = \sqrt{b^2 + 37^2}.$$

Yuck. We can relate a^2 and b^2 with $\sqrt{a^2 + 11^2} = \sqrt{b^2 + 37^2}$, but the expression $\sqrt{(a-b)^2 + (11-37)^2}$ stops us cold. Substituting for a or b looks pretty ugly.

Concept: When your initial approach to a problem hits a seemingly unsurpassable obstacle, take a step back and look for another way to view the problem that allows you to avoid the obstacle.

Using the lengths of the sides of the triangle didn't help, so let's consider the angles of the triangle. All three angles are 60 degrees, which might suggest the Law of Cosines. Unfortunately, that brings us back to the lengths of the sides, which gets us right back to the nasty expressions above.

Since the lengths of the sides are difficult to deal with in this problem, we'd like to find a useful fact about equilateral triangles that doesn't require that we use the actual lengths of the sides. Because $\triangle OAB$ is equilateral, point B is the image of rotating A by $\frac{\pi}{3}$ about the origin. Notice that we are not using the actual lengths of the sides here—we are only using the fact that the side lengths are the same. This is promising, so we investigate further. We have just learned a tool for investigating rotations about the origin—complex numbers.

To use complex numbers, we view the problem in the complex plane, so A is now the location of $a + 11i$ in the complex plane, and B corresponds to $b + 37i$. Because rotating A by $\frac{\pi}{3}$ degrees about the origin gives us point B, we must have

$$(a + 11i)e^{\pi i/3} = b + 37i \qquad \text{or} \qquad (a + 11i)e^{-\pi i/3} = b + 37i.$$

We'll work through the details of the first case, $(a + 11i)e^{\pi i/3} = b + 37i$. Expanding the left side gives

$$(a + 11i)e^{\pi i/3} = (a + 11i)\left(\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) = \frac{a}{2} - \frac{11\sqrt{3}}{2} + \frac{11}{2}i + \frac{a\sqrt{3}}{2}i,$$

so we have

$$\frac{a}{2} - \frac{11\sqrt{3}}{2} + \left(\frac{11}{2} + \frac{a\sqrt{3}}{2}\right)i = b + 37i.$$

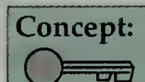
Equating the real parts and equating the imaginary parts of both sides gives

$$\frac{a}{2} - \frac{11\sqrt{3}}{2} = b,$$
$$\frac{11}{2} + \frac{a\sqrt{3}}{2} = 37.$$

The second equation gives us $a = 21\sqrt{3}$, and substituting this into the first equation gives $b = 5\sqrt{3}$, so $ab = 315$.

If we work through the details of the case $(a + 11i)e^{-\pi i/3} = b + 37i$ in the same manner as above, we'll find $a = -21\sqrt{3}$ and $b = -5\sqrt{3}$, so $ab = 315$ in this case, as well. This isn't a surprise: the equilateral triangle we produce in this case is the reflection over the imaginary axis of the one we found in the first case. □

In this problem, we combined our understanding of exponential and rectangular forms of complex numbers to tackle an analytic geometry problem that didn't appear to have anything to do with complex numbers. One of the clues that complex numbers might be helpful in this problem is:



Concept: Complex numbers can often be helpful in problems involving rotations, since rotations are easy to express with complex numbers.

Problem 8.24: Let ABC be a triangle in the complex plane whose vertices correspond to the complex numbers a, b, and c. Prove that $\triangle ABC$ is equilateral if and only if $a - b = \zeta(c - b)$, where ζ is a primitive sixth root of unity.

Solution for Problem 8.24: The two primitive sixth roots of unity are $e^{\pi i/3}$ and $e^{5\pi i/3}$, so we must show that $\triangle ABC$ is equilateral if and only if $a - b = e^{\pi i/3}(c - b)$ or $a - b = e^{5\pi i/3}(c - b)$. These equations look exactly like equations we encountered when discussing rotations. The equation $a - b = e^{\pi i/3}(c - b)$ tells us that A is the image of C upon a $\frac{\pi}{3}$ radian counterclockwise rotation about B. That is, we have $CB = BA$ and $\angle ABC = 60°$, from which we deduce that $\triangle ABC$ is equilateral. Similarly, the equation $a - b = e^{5\pi i/3}(c - b)$ means A is the image of C upon a $\frac{5\pi}{3}$ radian

counterclockwise rotation about B. Such a rotation is the same as a $2\pi - \frac{5\pi}{3} = \frac{\pi}{3}$ rotation *clockwise*, from which we again conclude that $CB = BA$ and $\angle ABC = 60°$, so $\triangle ABC$ is equilateral.

Conversely, if $\triangle ABC$ is equilateral, then a 60° rotation about B maps C to A, because $BC = BA$ and $\angle ABC = 60°$. If this rotation is counterclockwise, then we have $a - b = e^{\pi i/3}(c - b)$, and if it is clockwise, then $a - b = e^{5\pi i/3}(c - b)$.

Combining the previous two paragraphs, we conclude that $\triangle ABC$ is equilateral if and only if $a - b = \zeta(c - b)$, where ζ is a primitive sixth root of unity. □

Important: Let $\triangle ABC$ be a triangle in the complex plane whose vertices correspond to the complex numbers a, b, and c. Triangle ABC is equilateral if and only if $a - b = \zeta(c - b)$, where ζ is a primitive sixth root of unity.

Let's put our new tool for equilateral triangles to work.

Problem 8.25: Suppose $\triangle P_1P_2P_3$ and $\triangle Q_1Q_2Q_3$ are equilateral triangles with the same orientation. (By same orientation, we mean that the vertices are labeled in the same order, either clockwise or counterclockwise, in each triangle.) Let M_1, M_2, and M_3 be the midpoints of $\overline{P_1Q_1}$, $\overline{P_2Q_2}$, and $\overline{P_3Q_3}$, respectively. Show that $\triangle M_1M_2M_3$ is an equilateral triangle.

Solution for Problem 8.25: It's not clear where to start using the tools of Euclidean geometry, but with complex numbers, we don't even need a diagram to solve the problem quickly. We have a corresponding complex number p_i, q_i, or m_i for each of the vertices of the the three triangles, so that $m_i = (p_i + q_i)/2$ for $i = 1, 2, 3$, because M_i is the midpoint of $\overline{P_iQ_i}$.

Since $\triangle P_1P_2P_3$ is equilateral, we have

$$p_2 - p_1 = \zeta(p_3 - p_1)$$

for some primitive sixth root of unity ζ. Similarly, equilateral $\triangle Q_1Q_2Q_3$ gives us

$$q_2 - q_1 = \zeta(q_3 - q_1).$$

We know that the primitive sixth roots of unity in these two equations are the same because the triangles have the same orientation. If we didn't know that the two triangles have the same orientation, we couldn't take the next step, which is to add the two equations to produce expressions of the form $p_i + q_i$:

$$(p_2 + q_2) - (p_1 + q_1) = \zeta\left((p_3 + q_3) - (p_1 + q_1)\right).$$

Since $p_i + q_i = 2m_i$ for $i = 1, 2, 3$, we can write the equation above as $2m_2 - 2m_1 = \zeta(2m_3 - 2m_1)$. Dividing by 2 gives us $m_2 - m_1 = \zeta(m_3 - m_1)$, so $\triangle M_1M_2M_3$ is equilateral. □

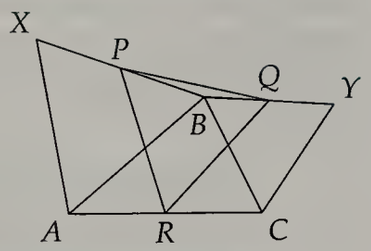
Problem 8.26: Points X and Y are outside $\triangle ABC$ such that $\triangle XAB$ and $\triangle YBC$ are equilateral and do not overlap $\triangle ABC$. Let P be the midpoint of $\overline{XB}$, Q be the midpoint of $\overline{YB}$, and R be the midpoint of $\overline{AC}$. Prove that $\triangle PQR$ is equilateral.

Solution for Problem 8.26: We know how to handle equilateral triangles and midpoints with complex numbers. As usual, we use the corresponding lowercase letter to represent the complex number that corresponds to each point. We'll present two solutions.

Solution 1: Work forwards. Our goal in this approach is to express p, q, and r in terms of a, b, and c, and then use these expressions to show that $\triangle PQR$ is equilateral. Since P and Q are the midpoints of $\overline{XB}$ and $\overline{YB}$, we'll have to find x and y in terms of a, b, and c first. We do so with equilateral triangles XAB and YBC. We focus on rotation about point B, the vertex the two triangles have in common. What's wrong with this start:

Bogus Solution: We have $x-b = \zeta(a-b)$ and $y-b = \zeta(c-b)$, where ζ is a primitive sixth root of unity. If we subtract the second equation from the first, we eliminate b.

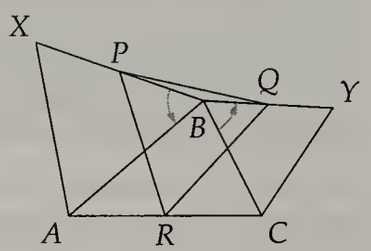
We see the problem with this start when we look at the diagram at right. The equation $x - b = \zeta(a - b)$ comes from the fact that X is the image of A upon a 60° rotation counterclockwise about B, but the equation $y - b = \zeta(c - b)$ comes from the fact that Y is the image of C upon a 60° rotation *clockwise* about B. The orientations of these two rotations are different, so we cannot use the same value of ζ for each rotation. We therefore cannot simply subtract the second equation from the first to eliminate b.



WARNING!! The orientations of rotations matter in complex numbers, just as they matter in geometry.

To fix this problem, we write equations using rotations of the same orientation, as marked in the diagram, so that we can use the same primitive sixth root of unity:

$$a - b = \zeta(x - b)$$
$$y - b = \zeta(c - b),$$



where ζ is a primitive sixth root of unity. Note that we cannot be sure which primitive sixth root of unity ζ is. In the diagram above, we would have $\zeta = e^{\pi i/3}$, but if the vertices of $\triangle ABC$ were labeled in the opposite order, we'd have $\zeta = e^{5\pi i/3}$. By simply writing ζ, we cover both cases, as long as we use the same orientation of rotation about B to produce the equations above.

Before continuing with finding p, q, and r, we see that we can simplify our equations for x and y a great deal by letting b be the origin, which gives $a = \zeta x$ and $y = \zeta c$. We therefore have

$$x = \frac{a}{\zeta}, \tag{8.5}$$
$$y = \zeta c. \tag{8.6}$$

Point B is also a natural choice for the origin because B is part of the definitions of both P and Q.

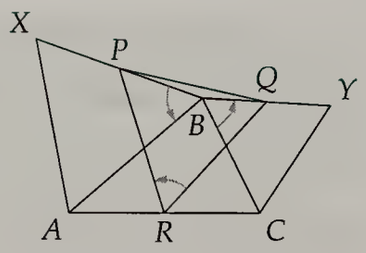
Concept: Strategically choosing the origin when applying complex numbers to geometry can greatly simplify the algebra you'll have to deal with.

Next, we find p, q, and r from the fact that P, Q, and R are the midpoints of $\overline{XB}$, $\overline{YB}$, and $\overline{AC}$:

$$p = \frac{x + b}{2} = \frac{x}{2},$$
$$q = \frac{y + b}{2} = \frac{y}{2},$$
$$r = \frac{a + c}{2}.$$

We have $x = \frac{a}{\zeta}$ from line (8.5) and $y = \zeta c$ from line (8.6), so we can write p and q in terms of a and c:

$$p = \frac{a}{2\zeta},$$
$$q = \frac{\zeta c}{2}.$$



We have a bit of a puzzle when we turn to $\triangle PQR$. We can test whether PQR is equilateral by testing rotation about any of the three vertices. Suppose we choose point R. Which should we try to confirm to check if $\triangle PQR$ is equilateral, $p-r = \zeta(q-r)$ or $q - r = \zeta(p - r)$? We choose $p - r = \zeta(q - r)$ because the orientation of 60° rotation about R that would map Q to P is the same as the orientation of the 60° rotation about B that maps X to A. This latter rotation is the one that gave us ζ, and we want to be able to use the same primitive sixth root of unity to avoid confusion and simplify the algebra. We have:

$$p - r = \frac{a}{2\zeta} - \frac{a+c}{2} = \frac{(1-\zeta)}{2\zeta}a - \frac{1}{2}c,$$
$$\zeta(q - r) = \zeta\left(\frac{\zeta c}{2} - \frac{a+c}{2}\right) = -\frac{\zeta}{2}a + \frac{\zeta^2 - \zeta}{2}c.$$

Our expressions for $p - r$ and $\zeta(q - r)$ don't appear to be the same. We can only be sure they're the same if the coefficients for a match and the coefficients for c match. In other words, we must check if

$$\frac{1-\zeta}{2\zeta} = -\frac{\zeta}{2} \qquad \text{(equating the coefficients of } a\text{)},$$
$$-\frac{1}{2} = \frac{\zeta^2 - \zeta}{2} \qquad \text{(equating the coefficients of } c\text{)}.$$

These both simplify to the same equation: $\zeta^2 - \zeta + 1 = 0$. So, if we can show that $\zeta^2 - \zeta + 1 = 0$ for both primitive sixth roots of unity ζ, then we can conclude that our expressions for $p - r$ and $\zeta(q - r)$ are the same. We could test that this equation is true for both primitive sixth roots ζ, but there's a slicker way to confirm that $\zeta^2 - \zeta + 1 = 0$. First, we note that ζ is a sixth root of unity, so it is a solution to $z^6 - 1 = 0$. Factoring the left side as a difference of squares gives

$$(z^3 - 1)(z^3 + 1) = 0.$$

The roots of $z^3 - 1$ are cube roots of unity, so they cannot be primitive sixth roots. Therefore, ζ is a root of $z^3 + 1$. Factoring this as a sum of cubes gives $(z + 1)(z^2 - z + 1)$. Since -1 is a primitive square root of unity rather than a primitive sixth root of unity, we know that ζ is not a root of $z + 1$. So, we know that the two primitive sixth roots of unity are roots of $z^2 - z + 1 = 0$. This means we have $\zeta^2 - \zeta + 1 = 0$, as desired. Therefore, we have $1 - \zeta = -\zeta^2$ and $\zeta^2 - \zeta = -1$, and we can write our expressions for $p - r$ and $\zeta(q - r)$ as follows:

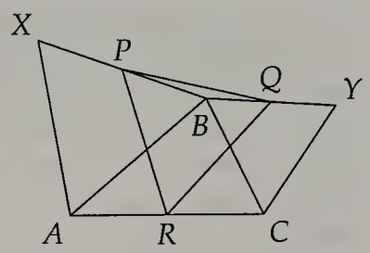
$$p - r = \frac{(1-\zeta)}{2\zeta}a - \frac{1}{2}c = \frac{-\zeta^2}{2\zeta}a - \frac{1}{2}c = -\frac{\zeta}{2}a - \frac{1}{2}c,$$
$$\zeta(q - r) = -\frac{\zeta}{2}a + \frac{\zeta^2 - \zeta}{2}c = -\frac{\zeta}{2}a - \frac{1}{2}c.$$

Therefore, we have $p - r = \zeta(q - r)$, as desired.

One key step in this solution was noticing that two very different-looking expressions involving ζ are equivalent.

Concept: When working with roots of unity, we will occasionally encounter two very different-looking expressions that we can show are equivalent.

Looking back, there wasn't any reason to know immediately that $\frac{1-\zeta}{2\zeta} = -\frac{\zeta}{2}$, but once we discovered that we needed this to be true to solve the problem, we focused on it and were able to prove it. What if we had taken this strategy for the whole problem—what if we focused first on what we wanted to be true?



Solution 2: Work backwards. In this solution, we start with expressions for p, q, and r in terms of x, y, a, b, and c, and work backwards from an equation that implies that $\triangle PQR$ is equilateral. Our goal is to find an equivalent equation that we can prove is true.

As in the first solution we find expressions for p, q, and r using the fact that P, Q, and R are the midpoints of $\overline{XB}$, $\overline{YB}$, and $\overline{AC}$:

$$p = \frac{x+b}{2},$$
$$q = \frac{y+b}{2},$$
$$r = \frac{a+c}{2}.$$

There are a variety of equations we can write in terms of p, q, and r that are true if and only if $\triangle PQR$ is equilateral. We choose one involving $p-q$ since this will eliminate b when we substitute the expressions above. Specifically, if we can show that $p-q = \zeta(r-q)$, where ζ is a primitive sixth root of unity, then we can deduce that $\triangle PQR$ is equilateral. Substituting our expressions for p, q, and r, we wish to show that

$$\frac{x-y}{2} = \zeta\left(\frac{a+c-y-b}{2}\right). \tag{8.7}$$

We turn to equilateral triangles $\triangle XAB$ and $\triangle YBC$ in order to produce equations involving x, y, a, b, and c. In writing equations including ζ, we must make sure to use the same orientation of rotation that we used when writing $p-q = \zeta(r-q)$.

There are three different equations we could write for each triangle, one for rotation about each vertex. Rather than write them all down, we use our target Equation (8.7) as a guide. Since $\overline{AB}$ is a side of $\triangle XAB$ and $\overline{YC}$ is a side of $\triangle YBC$, we can write the right side of Equation (8.7) in terms of complex numbers corresponding to sides of these two triangles:

$$\frac{x-y}{2} = \zeta\left(\frac{(a-b)+(c-y)}{2}\right).$$

The $a-b$ tells us to focus on the rotation about B in $\triangle XAB$ and the $c-y$ tells us to focus on the rotation about Y in $\triangle YBC$. These rotations (taking the same orientation we use to write $p-q = \zeta(r-q)$) give us

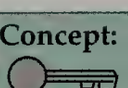
$$x-b = \zeta(a-b),$$
$$b-y = \zeta(c-y).$$

Aha! If we simply add these two equations, we get

$$x-y = \zeta(a+c-y-b).$$

Dividing this equation by 2 gives us Equation (8.7), which we have shown is equivalent to $p-q = \zeta(r-q)$. Therefore, we can deduce that $\triangle PQR$ is equilateral. □

Our second solution is much simpler than the first. If we had been lucky in choosing the right equations for $\triangle XAB$ and $\triangle YBC$ in the first solution, we might have stumbled on the simpler approach we found in the second solution. The key to our second solution was working backwards to give us clues how to approach the problem in a way that would provide the desired result quickly.



Concept: Working backwards from the desired result is a powerful problem solving strategy.

Equilateral triangles and squares are not the only regular polygons for which complex numbers can be helpful.

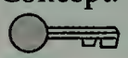
Problem 8.27: Let $n \geq 3$ be a positive integer, and let $P_0P_1 \cdots P_{n-1}$ be a regular n-gon inscribed in a circle with radius 1. Compute $P_0P_1 \cdot P_0P_2 \cdots P_0P_{n-1}$.

Solution for Problem 8.27: If "circle with radius 1" isn't enough of a hint to try complex numbers, "regular n-gon" sure is, because we have an easy way to represent a regular n-gon on the complex plane—with the n^{th} roots of unity. We let $p_0, p_1, \ldots, p_{n-1}$ be the n^{th} roots of unity, and let P_k correspond to p_k for each root. Moreover, we let $p_k = e^{2\pi ik/n}$ for each k, which makes $p_0 = 1$. This is particularly convenient because we can now express each length P_0P_k as $|p_0 - p_k| = |1 - e^{2\pi ik/n}|$. Therefore, the desired product is

$$\left|1 - e^{1\cdot 2\pi i/n}\right| \cdot \left|1 - e^{2\cdot 2\pi i/n}\right| \cdot \left|1 - e^{3\cdot 2\pi i/n}\right| \cdots \left|1 - e^{(n-1)\cdot 2\pi i/n}\right|.$$

If we don't recognize what to do right away, we can experiment.

Concept: When a problem is given in terms of an integer n, experimenting with small values of n can give clues to the solution.



When $n = 2$, the product is simply $|1 - (-1)| = 2$. When $n = 3$, we have $|1 - \omega| \cdot |1 - \omega^2|$, where $\omega = e^{2\pi i/3}$. Computing this product gives

$$|1 - \omega| \cdot \left|1 - \omega^2\right| = \left|(1 - \omega)(1 - \omega^2)\right| = \left|1 - \omega^2 - \omega + \omega^3\right|.$$

We can simplify this expression by noting that because ω is a primitive cube root of unity, we have $\omega^3 - 1 = 0$. Therefore, not only do we have $\omega^3 = 1$, but factoring gives us $(\omega - 1)(\omega^2 + \omega + 1) = 0$. Since $\omega \neq 1$, we have $\omega^2 + \omega + 1 = 0$, so $\omega^2 + \omega = -1$. Therefore, we have

$$\left|1 - \omega^2 - \omega + \omega^3\right| = |1 - (\omega^2 + \omega) + 1| = 3.$$

Important: If ω is a primitive cube root of unity, then $\omega^2 + \omega + 1 = 0$.

So, when $n = 2$, the product is 2, and when $n = 3$, the product is 3. We now have a guess for what we'll get when $n = 4$. Let's check:

$$|1 - i| \cdot |1 - (-1)| \cdot |1 - (-i)| = 2|(1 - i)(1 + i)| = 4.$$

We now have a pretty strong suspicion that the product equals n for any positive integer n, but it's still not so clear how to prove it. Let's take another look at the product, and think about where else we have seen similar products:

$$\left|1 - e^{1\cdot 2\pi i/n}\right| \cdot \left|1 - e^{2\cdot 2\pi i/n}\right| \cdot \left|1 - e^{3\cdot 2\pi i/n}\right| \cdots \left|1 - e^{(n-1)\cdot 2\pi i/n}\right|.$$

Each term is the difference of 1 and an n^{th} root of unity. We have seen a product of differences involving the n^{th} roots of unity before, in the factorization of the polynomial $z^n - 1$. If $f(z) = z^n - 1$, then we can write $f(z)$ as

$$f(z) = (z - 1)\left(z - e^{1\cdot 2\pi i/n}\right)\left(z - e^{2\cdot 2\pi i/n}\right)\left(z - e^{3\cdot 2\pi i/n}\right) \cdots \left(z - e^{(n-1)\cdot 2\pi i/n}\right).$$

Except for the first term, $z - 1$, this looks almost exactly like our desired product. Fortunately, we can also factor

$f(z)$ as $(z-1)(z^{n-1}+z^{n-2}+z^{n-3}+\cdots+1)$, which allows us to divide by $z-1$ to get

$$\begin{aligned}\left(z-e^{1\cdot 2\pi i/n}\right)\left(z-e^{2\cdot 2\pi i/n}\right)\left(z-e^{3\cdot 2\pi i/n}\right)\cdots\left(z-e^{(n-1)\cdot 2\pi i/n}\right) &= \frac{f(z)}{z-1}\\ &= \frac{(z-1)(z^{n-1}+z^{n-2}+z^{n-3}+\cdots+1)}{z-1}\\ &= z^{n-1}+z^{n-2}+z^{n-3}+\cdots+1.\end{aligned}$$

We also could have jumped straight to

$$\left(z-e^{1\cdot 2\pi i/n}\right)\cdot\left(z-e^{2\cdot 2\pi i/n}\right)\left(z-e^{3\cdot 2\pi i/n}\right)\cdot\cdots\cdot\left(z-e^{(n-1)\cdot 2\pi i/n}\right)=z^{n-1}+z^{n-2}+z^{n-3}+\cdots+1$$

by noting that the roots of the polynomial on the right are all the n^{th} roots of unity except 1. Speaking of 1, we let $z=1$ in this equation to find that

$$\begin{aligned}\left|1-e^{1\cdot 2\pi i/n}\right|\cdot\left|1-e^{2\cdot 2\pi i/n}\right|\cdot\cdots\cdot\left|1-e^{(n-1)\cdot 2\pi i/n}\right| &= \left|\left(1-e^{1\cdot 2\pi i/n}\right)\cdot\left(1-e^{2\cdot 2\pi i/n}\right)\cdot\cdots\cdot\left(1-e^{(n-1)\cdot 2\pi i/n}\right)\right|\\ &= \left|1^{n-1}+1^{n-2}+1^{n-3}+\cdots+1^1+1\right|\\ &= n,\end{aligned}$$

as expected. □

Exercises

8.4.1 Suppose $PQRS$ is a square in the complex plane, and that p and q correspond to P and Q, respectively. Find the two possible complex numbers that correspond to R in terms of p and q.

8.4.2 Let O be the center of a circle ω. Points A, B, C, D, E, F on ω are chosen such that the triangles OAB, OCD, OEF are equilateral and have the same orientation. Let L, M, N be the midpoints of $\overline{BC}$, $\overline{DE}$, $\overline{FA}$, respectively. Prove that triangle LMN is equilateral.

8.4.3 Let $ABCD$ be a parallelogram. Suppose we construct equilateral triangles ABG and BCF externally on sides $\overline{AB}$ and $\overline{BC}$, respectively. Show that triangle DFG is equilateral.

8.4.4 Prove that each of the following are necessary and sufficient conditions for the triangle formed by plotting complex numbers a, b, and c in the complex plane to be an equilateral triangle. (That is, show that if a, b, c satisfy the condition, then the triangle is equilateral, and show that if the triangle is equilateral, then a, b, and c satisfy the given condition.)

(a) $a-b=\omega(b-c)$, where ω is a primitive cube root of unity. **Hints:** 281

(b) $a+\omega b+\omega^2 c=0$, where ω is a primitive cube root of unity. **Hints:** 14

8.4.5⋆ Inside square $ABCD$, we construct equilateral triangles ABK, BCL, CDM, and DAN. Prove that the 12 points which are the midpoints of $\overline{KL}$, $\overline{LM}$, $\overline{MN}$, $\overline{NK}$, $\overline{AK}$, $\overline{BK}$, $\overline{BL}$, $\overline{CL}$, $\overline{CM}$, $\overline{DM}$, $\overline{DN}$, amd $\overline{AN}$ form the vertices of a regular dodecagon.

8.5⋆ Classic Theorems

In this section, we exhibit the power of complex numbers by proving four classic geometry theorems. Some of these proofs are quite difficult, so don't feel bad if you can't find them on your own. But, do work through and enjoy them!

Problems

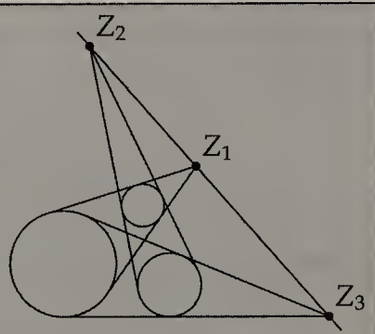
Problem 8.28: Ptolemy's Inequality: Prove that for any four points A, B, C, D in a plane, we have

$$(AB)(CD) + (BC)(AD) \geq (AC)(BD).$$

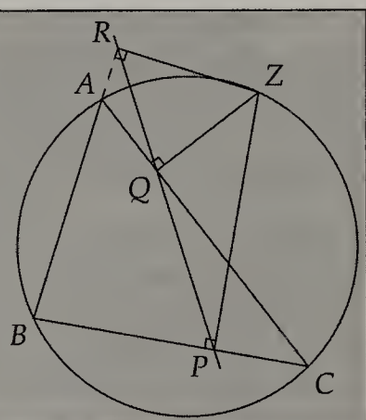
Hints: 56

Problem 8.29: Outer Napoleon Triangle: Starting with $\triangle PQR$, we construct equilateral triangles $\triangle APQ$, $\triangle BQR$, and $\triangle CRP$ such that these triangles do not overlap $\triangle PQR$. Prove that the triangle formed by connecting the centers of these three equilateral triangles is itself equilateral. **Hints:** 243

Problem 8.30: Monge's Theorem: Three circles are drawn in a plane with different radii such that no circle is inside another. The common external tangents to each pair of circles are drawn, as shown. The intersections of pairs of common external tangents are labeled Z_1, Z_2, and Z_3, as shown. Show that Z_1, Z_2, and Z_3 are collinear. **Hints:** 218, 284



Problem 8.31: Simson Line: Let Z be a point on the circumcircle of $\triangle ABC$ besides A, B, and C. Show that the feet of the perpendiculars from Z to the lines containing the sides of $\triangle ABC$ are collinear. **Hints:** 120, 61



We used a geometric explanation of the Triangle Inequality to develop an inequality relating complex numbers. Let's try turning the tables, and use complex numbers to develop a geometric inequality.

Problem 8.28: Ptolemy's Inequality: Show that for any four points A, B, C, D in a plane, we have

$$(AB)(CD) + (BC)(AD) \geq (AC)(BD).$$

Solution for Problem 8.28: We let complex numbers a, b, c, d correspond to A, B, C, D, respectively, so that

$$(AB)(CD) + (BC)(AD) = |b-a| \cdot |d-c| + |c-b| \cdot |d-a|.$$

We have to prove an inequality, and the Triangle Inequality gives us a tool to produce inequalities regarding magnitudes of complex numbers. To apply $|w| + |z| \geq |w+z|$, we write

$$|b-a| \cdot |d-c| + |c-b| \cdot |d-a| = |(b-a)(d-c)| + |(c-b)(d-a)|.$$

Applying the Triangle Inequality gives

$$|(b-a)(d-c)| + |(c-b)(d-a)| \geq |(b-a)(d-c) + (c-b)(d-a)|.$$

(Note that our expression for $(AB)(CD) + (BC)(AD)$ is on the larger side of this inequality, which gives us some confidence that we're moving in the correct direction.) We can see that some of the terms in the expression on the lesser side will cancel when we expand, so we give that a try:

$$\begin{aligned} |(b-a)(d-c) + (c-b)(d-a)| &= |bd - bc - ad + ac + cd - ac - bd + ab| \\ &= |ab + cd - bc - ad|. \end{aligned}$$

If we can't see that this expression factors, then we can always get a hint from the desired inequality. We want to show that $(AB)(CD) + (BC)(AD)$ is at least $(AC)(BD)$, which equals $|(c-a)(d-b)|$. Expanding this gives us exactly the expression above, $|ab + cd - bc - ad|$, and now we have our proof:

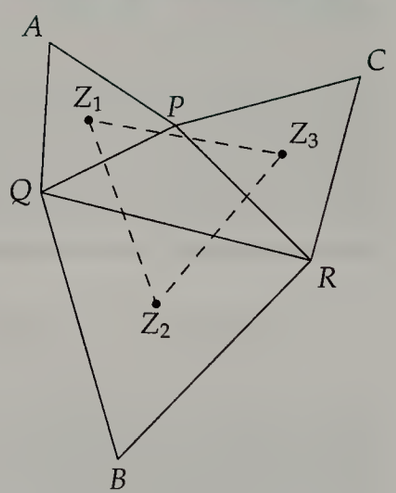
$$\begin{aligned} (AB)(CD) + (BC)(AD) &= |b-a| \cdot |d-c| + |c-b| \cdot |d-a| \\ &= |(b-a)(d-c)| + |(c-b)(d-a)| \\ &\geq |(b-a)(d-c) + (c-b)(d-a)| \\ &= |ab + cd - bc - ad| \\ &= |(c-a)(d-b)| \\ &= (AC)(BD). \end{aligned}$$

□

As an extra challenge, see if you can discover the conditions under which equality holds in Ptolemy's Inequality.

Problem 8.29: Outer Napoleon Triangle: Starting with $\triangle PQR$, we construct triangles equilateral triangles $\triangle APQ$, $\triangle BQR$, and $\triangle CRP$ such that these triangles do not overlap $\triangle PQR$. Prove that the triangle formed by connecting the centers of these three equilateral triangles is equilateral.

Solution for Problem 8.29: As usual, we let the lowercase version of each point be the complex number corresponding to that point. We start by finding expressions for a, b, and c in terms of p, q, and r. From equilateral triangle APQ, we have $a - q = \zeta(p - q)$, where ζ is a primitive sixth root of unity. Therefore, we have $a = \zeta(p - q) + q$. Similarly, from triangles BQR and CRP, we have $b = \zeta(q - r) + r$ and $c = \zeta(r - p) + p$.



Note that we are careful to use the same orientation for each rotation so that we can use the same primitive sixth root. For example, in $\triangle APQ$, we get $a - q = \zeta(p - q)$ from considering the $\frac{\pi}{3}$ counterclockwise rotation of P about Q in the diagram at right, and we likewise consider counterclockwise rotations to produce our expressions for b and c.

The medians of an equilateral triangle intersect at the center of the triangle. As you should have showed in Exercise 8.2.5, the centroid of a triangle in the complex plane is the average of the vertices of the triangle. So, if we let Z_1, Z_2, and Z_3 be the centers of triangles APQ, BQR, and CRP, respectively, we have

$$\begin{aligned} z_1 &= \frac{a+p+q}{3} = \frac{\zeta(p-q)+q+p+q}{3} = \frac{(1+\zeta)p + (2-\zeta)q}{3}, \\ z_2 &= \frac{b+q+r}{3} = \frac{\zeta(q-r)+r+q+r}{3} = \frac{(1+\zeta)q + (2-\zeta)r}{3}, \\ z_3 &= \frac{c+r+p}{3} = \frac{\zeta(r-p)+p+r+p}{3} = \frac{(1+\zeta)r + (2-\zeta)p}{3}. \end{aligned}$$

We can prove that triangle $Z_1Z_2Z_3$ is equilateral by showing that $(z_3 - z_1) = \zeta(z_2 - z_1)$. (Note that we are careful again to use the same orientation of rotation as with the earlier equilateral triangles.) Substituting, we find that

$$z_3 - z_1 = \frac{(1+\zeta)r + (2-\zeta)p}{3} - \frac{(1+\zeta)p + (2-\zeta)q}{3} = \frac{(1-2\zeta)p + (-2+\zeta)q + (1+\zeta)r}{3},$$

and

$$z_2 - z_1 = \frac{(1+\zeta)q + (2-\zeta)r}{3} - \frac{(1+\zeta)p + (2-\zeta)q}{3} = \frac{(-1-\zeta)p + (-1+2\zeta)q + (2-\zeta)r}{3},$$

so

$$\zeta(z_2 - z_1) = \frac{(-\zeta - \zeta^2)p + (-\zeta + 2\zeta^2)q + (2\zeta - \zeta^2)r}{3}.$$

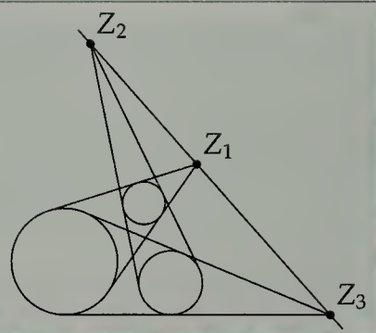
Uh-oh; our expression for $\zeta(z_2 - z_1)$ doesn't seem to match our expression for $z_3 - z_1$. However, we do have ζ^2 in the expression for $\zeta(z_2 - z_1)$, and back on page 283 we saw a convenient manipulation for primitive sixth roots of unity. Specifically, if ζ is a primitive sixth root of unity, then $\zeta^2 - \zeta + 1 = 0$, so $\zeta^2 = \zeta - 1$. Making this substitution in our expression for $\zeta(z_2 - z_1)$ gives

$$\zeta(z_2 - z_1) = \frac{(-\zeta - (\zeta-1))p + (-\zeta + 2(\zeta-1))q + (2\zeta - (\zeta-1))r}{3} = \frac{(1-2\zeta)p + (-2+\zeta)q + (1+\zeta)r}{3},$$

which does indeed match our expression for $z_3 - z_1$. Since $z_3 - z_1 = \zeta(z_2 - z_1)$, we know that $Z_1Z_2Z_3$ is equilateral. □

Sidenote: The "Outer Napoleon Triangle" is equilateral; is the "inner" one also equilateral? If you construct equilateral triangles $\triangle APQ$, $\triangle BQR$, and $\triangle CRP$ so that they overlap $\triangle PQR$, is the triangle formed by connecting the centers of these three equilateral triangles also equilateral?

Problem 8.30: Monge's Theorem: Three circles are drawn in a plane with different radii such that no circle is inside another. The common external tangents to each pair of circles are drawn, as shown. The intersections of pairs of common external tangents are labeled Z_1, Z_2, and Z_3, as shown. Show that Z_1, Z_2, and Z_3 are collinear.



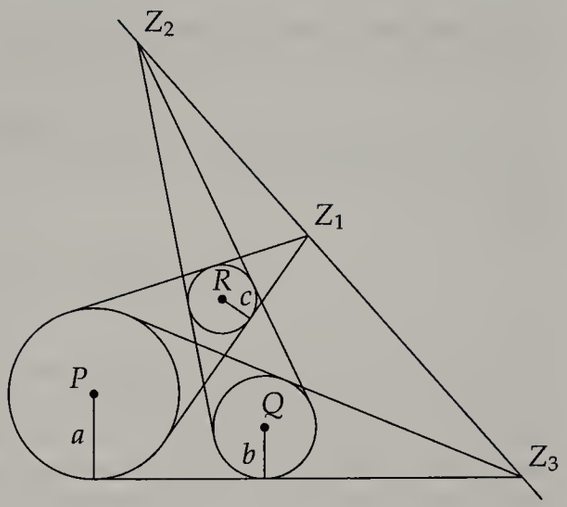
Solution for Problem 8.30: The only clue that we might be able to solve this problem with complex numbers is that we wish to prove three points collinear. We have a pretty straightforward condition for collinearity in the complex plane. Letting z_1, z_2, and z_3 be complex numbers corresponding to points Z_1, Z_2, and Z_3, the points are collinear if and only if

$$(z_1 - z_2)(\overline{z_3} - \overline{z_2}) - (\overline{z_1} - \overline{z_2})(z_3 - z_2) = 0.$$

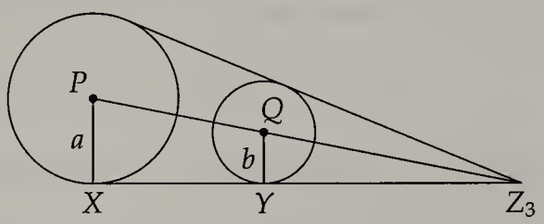
Extra! There's a wonderful proof of Monge's Theorem that involves thinking of the problem as a cross-section of a three-dimensional configuration. As a hint, consider the centers of the three circles as the centers of three spheres whose radii equal the radii of the corresponding circles.

Solution on page 297.

Now our goal is to find z_1, z_2, and z_3 in terms of the radii and the complex numbers corresponding to the centers of the circles. We let P, Q, and R be the centers of the circles, and we'll refer to the circles as $\odot P$, $\odot Q$, and $\odot R$. Let $\odot P$, $\odot Q$, and $\odot R$ have radii a, b, and c, respectively, with $a > b > c$. The two common external tangents to $\odot P$ and $\odot Q$ meet at Z_3, the two common external tangents to $\odot Q$ and $\odot R$ meet at Z_2, and the two common external tangents to $\odot P$ and $\odot R$ meet at Z_1.



We start with Z_3. P and Q are equidistant from the common external tangents to $\odot P$ and $\odot Q$, so they are on the angle bisector of the angle formed by these tangents, which means P, Q, and Z_3 are collinear, as shown. We draw radii $\overline{PX}$ and $\overline{QY}$ to the common tangent, thereby forming similar right triangles PXZ_3 and QYZ_3.



Since P, Q, and Z_3 are collinear, we know that $(z_3 - p)/(q - p)$ is real and equals the ratio of magnitudes of $z_3 - p$ and $q - p$. That is, we have

$$\frac{z_3 - p}{q - p} = \frac{PZ_3}{PQ}.$$

We can use similar triangles PXZ_3 and QYZ_3 to express PZ_3/PQ in terms of the radii of the circles. Specifically, we have $QZ_3/PZ_3 = b/a$, so $PQ/PZ_3 = (PZ_3 - QZ_3)/PZ_3 = 1 - (b/a) = (a - b)/a$, which means

$$\frac{z_3 - p}{q - p} = \frac{PZ_3}{PQ} = \frac{a}{a - b}.$$

Therefore, we have

$$z_3 = p + \frac{a}{a - b}(q - p) = \frac{ap - bp}{a - b} + \frac{aq - ap}{a - b} = \frac{aq - bp}{a - b}.$$

Similarly, we have

$$z_2 = \frac{br - cq}{b - c},$$
$$z_1 = \frac{ar - cp}{a - c}.$$

We have to compute the expression

$$(z_1 - z_2)(\overline{z_3} - \overline{z_2}) - (\overline{z_1} - \overline{z_2})(z_3 - z_2),$$

since we wish to show that it equals 0. This looks pretty scary, until we realize that we haven't chosen our origin yet. If we choose z_2 to be the origin, then this expression becomes a much more manageable

$$z_1\overline{z_3} - \overline{z_1}z_3.$$

Moreover, by letting z_2 be the origin, the expression above for z_2 must be 0, so

$$\frac{br - cq}{b - c} = 0,$$

which gives us $r = \frac{cq}{b}$, thereby allowing us to eliminate r from our expressions for z_1 and z_3:

$$z_1 = \frac{ar - cp}{a - c} = \frac{\frac{acq}{b} - cp}{a - c} = \frac{acq - bcp}{b(a - c)} = \frac{c}{b(a - c)}(aq - bp),$$
$$z_3 = \frac{aq - bp}{a - b} = \frac{1}{a - b}(aq - bp).$$

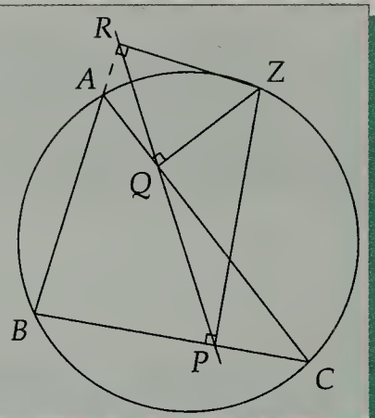
Concept: Don't always specify the origin right away—sometimes waiting will allow you to exploit symmetry in a problem or will reveal a surprisingly useful choice for the origin.

This particularly useful choice for the origin pretty much finishes off the problem. We now see that we don't even have to compute $z_1\overline{z_3} - \overline{z_1}z_3$. We have

$$\frac{z_1}{z_3} = \frac{\frac{c}{b(a-c)}(aq - bp)}{\frac{1}{a-b}(aq - bp)} = \frac{c(a - b)}{b(a - c)},$$

which is real because a, b, and c are real. Since z_1/z_3 is a real number, the origin, Z_1, and Z_3 are collinear. But Z_2 is the origin, so we have shown that Z_1, Z_2, and Z_3 are collinear! (Moreover, we have also proved a relationship among Z_1Z_2/Z_3Z_2 and the radii of the circles.) □

Problem 8.31: Simson Line: Let Z be a point on the circumcircle of $\triangle ABC$ besides A, B, and C. Show that the feet of the perpendiculars from Z to the lines containing the sides of $\triangle ABC$ are collinear.



Solution for Problem 8.31: Let the feet of the perpendiculars from Z to $\overleftrightarrow{BC}$, $\overleftrightarrow{AC}$, and $\overleftrightarrow{AB}$ be P, Q, and R, respectively. As usual, we let the lowercase letters of each of the point labels be the corresponding complex numbers. In order to show that P, Q, and R are collinear, we would like to show that

$$(p - q)(\overline{p} - \overline{r}) - (\overline{p} - \overline{q})(p - r) = 0.$$

Our first step is to find p, q, and r in terms of a, b, c, and z. Fortunately, we only have to find one of them, since they'll all have the same form. Once we have one of them, we hopefully can use that form to determine the other two.

Since P is on $\overleftrightarrow{BC}$, complex numbers p, b, and c are collinear in the complex plane. This means that

$$(p - b)(\overline{b} - \overline{c}) - (\overline{p} - \overline{b})(b - c) = 0. \tag{8.8}$$

We still need to find another equation for p (the one we just found only comes from the fact that P is on $\overleftrightarrow{BC}$, which isn't enough to uniquely specify P). We turn to our other restriction on P, which is that $\overleftrightarrow{ZP} \perp \overleftrightarrow{BC}$. This means that we have

$$(p - z)(\overline{b} - \overline{c}) + (\overline{p} - \overline{z})(b - c) = 0. \tag{8.9}$$

Comparing this to Equation (8.8), we see something promising. If we add the two equations, we cancel $\overline{p}$, and we can solve for p. So, we add Equations (8.8) and (8.9) by combining $(\overline{b}-\overline{c})$ terms and $(b-c)$ terms:

$$((p-b)+(p-z))(\overline{b}-\overline{c})+(-(\overline{p}-\overline{b})+(\overline{p}-\overline{z}))(b-c)=0.$$

We therefore have $(2p-b-z)(\overline{b}-\overline{c})+(\overline{b}-\overline{z})(b-c)=0$, and solving for p gives

$$p=\frac{1}{2}\left(-\frac{(\overline{b}-\overline{z})(b-c)}{\overline{b}-\overline{c}}+b+z\right). \tag{8.10}$$

(Note that we can divide by $\overline{b}-\overline{c}$ because $b\neq c$.) Well, at least we found p, but that expression on the right is quite a mess. We're not stuck yet, though. We haven't chosen our origin. One natural choice might appear to be z, but then we're left with $\overline{b}$ and $\overline{c}$, and we'd prefer to get rid of them, too. Otherwise, after writing similar expressions for q and r, we'll have $\overline{a}$, $\overline{b}$, and $\overline{c}$ to deal with on top of a, b, and c. (Try it and see—you'll get a nice enough expression for p, but once you join it with the expressions for q and r, you'll have quite an intimidating mess on your hands for $(p-q)(\overline{p}-\overline{r})-(\overline{p}-\overline{q})(p-r)$.)

So, we go back to the drawing board on choosing the origin. We focus on the conjugates, since they scared us off the choice of z as the origin. A natural choice that presents itself for the origin is the center of the circle. By choosing the center of the circle as the origin and letting the radius of the circle be 1, we have $|a|=|b|=|c|=|z|=1$, so $a\overline{a}=b\overline{b}=c\overline{c}=z\overline{z}=1$. This allows us to get rid of all of the conjugates in our expression for p in Equation (8.10)! Let's see what happens when we substitute to get rid of the conjugates:

$$p=\frac{1}{2}\left(-\frac{\left(\frac{1}{b}-\frac{1}{z}\right)(b-c)}{\frac{1}{b}-\frac{1}{c}}+b+z\right)=\frac{1}{2}\left(-\frac{\left(\frac{z-b}{bz}\right)(b-c)}{\frac{c-b}{bc}}+b+z\right)=\frac{1}{2}\left(\frac{(z-b)c}{z}+b+z\right)=\frac{1}{2}\left(b+c+z-\frac{bc}{z}\right).$$

Now that's something we can work with! By symmetry, we can now write expressions for q and r, and we have

$$\begin{aligned}p&=\frac{1}{2}\left(b+c+z-\frac{bc}{z}\right),\\ q&=\frac{1}{2}\left(c+a+z-\frac{ca}{z}\right),\\ r&=\frac{1}{2}\left(a+b+z-\frac{ab}{z}\right).\end{aligned}$$

Now we're ready to tackle $(p-q)(\overline{p}-\overline{r})-(\overline{p}-\overline{q})(p-r)$. If we can show this equals 0, then we know that P, Q, and R are collinear. Each of the differences in this expression gives us a lot of cancellation, so it's not nearly as scary as it looks. We'll start by simplifying $(p-q)(\overline{p}-\overline{r})$, since we can simply take the conjugate of the result to get $(\overline{p}-\overline{q})(p-r)$. We find

$$\begin{aligned}(p-q)(\overline{p}-\overline{r})&=\frac{1}{4}\left(b-a-\frac{bc}{z}+\frac{ac}{z}\right)\left(\overline{c}-\overline{a}-\frac{\overline{b}\overline{c}}{\overline{z}}+\frac{\overline{a}\overline{b}}{\overline{z}}\right)\\ &=\frac{1}{4}\left(b-a-(b-a)\frac{c}{z}\right)\left(\overline{c}-\overline{a}-(\overline{c}-\overline{a})\frac{\overline{b}}{\overline{z}}\right)\\ &=\frac{1}{4}(b-a)\left(1-\frac{c}{z}\right)(\overline{c}-\overline{a})\left(1-\frac{\overline{b}}{\overline{z}}\right)\\ &=\frac{1}{4}(b-a)\left(\frac{z-c}{z}\right)(\overline{c}-\overline{a})\left(\frac{\overline{z}-\overline{b}}{\overline{z}}\right)\\ &=\frac{1}{4z\overline{z}}(b-a)(z-c)(\overline{c}-\overline{a})(\overline{z}-\overline{b}).\end{aligned}$$

We're very close now! We can again substitute for the conjugates (so $z\bar{z} = 1$, for example), and we have

$$(p-q)(\bar{p}-\bar{r}) = \frac{1}{4}(b-a)(z-c)\left(\frac{1}{c}-\frac{1}{a}\right)\left(\frac{1}{z}-\frac{1}{b}\right) = \frac{1}{4abcd}(b-a)(z-c)(a-c)(b-z).$$

Taking the conjugate of this, and then substituting for $\bar{a}$, $\bar{b}$, $\bar{c}$, and $\bar{z}$, gives

$$\begin{aligned}(\bar{p}-\bar{q})(p-r) &= \frac{1}{4\bar{a}\bar{b}\bar{c}\bar{d}}(\bar{b}-\bar{a})(\bar{z}-\bar{c})(\bar{a}-\bar{c})(\bar{b}-\bar{z}) \\ &= \frac{abcd}{4}\left(\frac{1}{b}-\frac{1}{a}\right)\left(\frac{1}{z}-\frac{1}{c}\right)\left(\frac{1}{a}-\frac{1}{c}\right)\left(\frac{1}{b}-\frac{1}{z}\right) \\ &= \frac{abcd}{4}\left(\frac{a-b}{ab}\right)\left(\frac{c-z}{cz}\right)\left(\frac{c-a}{ac}\right)\left(\frac{z-b}{bz}\right) \\ &= \frac{1}{4abcd}(a-b)(c-z)(c-a)(z-b).\end{aligned}$$

This equals our expression for $(p-q)(\bar{p}-\bar{r})$ because each of the four differences here is the opposite of the corresponding difference in our expression for $(p-q)(\bar{p}-\bar{r})$. Therefore, $(p-q)(\bar{p}-\bar{r})$ equals its own conjugate, which means it is real. This tells us that P, Q, and R are collinear. The line that contains these points is called a **Simson Line** of the triangle.□

The key step in this solution can be a very powerful strategy:

Concept: When applying complex numbers to a geometry problem involving a single circle, letting the circle be the unit circle centered at the origin can be a very powerful strategy. One reason this is so useful is that for any point z on the circle, we have $\bar{z} = \frac{1}{z}$.

8.6 Summary

The image of a translation of z by the complex number w is $z + w$, the image of a dilation of z about the origin by scale factor c is cz, and the image of a rotation of z about a point w counterclockwise by an angle θ is $e^{i\theta}(z-w)+w$.

Important: The midpoint of the segment connecting z_1 and z_2 in the complex plane is $\frac{z_1+z_2}{2}$.

Important: Let A, B, C, and D be four different points in the complex plane, and let a, b, c, and d be complex numbers corresponding to these points.

- We have $\overleftrightarrow{AB} \parallel \overleftrightarrow{CD}$ if and only if $(b-a)(\bar{d}-\bar{c}) - (\bar{b}-\bar{a})(d-c) = 0$. This is equivalent to the condition that $(b-a)/(d-c)$ is real.
- We have $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$ if and only if $(b-a)(\bar{d}-\bar{c}) + (\bar{b}-\bar{a})(d-c) = 0$. This is equivalent to the condition that $(b-a)/(d-c)$ is imaginary.

Important: If complex numbers w and z correspond to points W and Z in the complex plane, then $WZ^2 = (w-z)(\bar{w}-\bar{z})$.

Important:

For any complex numbers $z_1, z_2, z_3, \ldots, z_n$, we have

$$|z_1| + |z_2| + \cdots + |z_n| \geq |z_1 + z_2 + \cdots + z_n|,$$

where equality holds if and only if $z_1, z_2, z_3, \ldots, z_n$ are on the same ray from the origin.

Important:

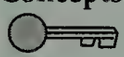
Let $\triangle ABC$ be a triangle in the complex plane whose vertices correspond to the complex numbers a, b, and c. Triangle ABC is equilateral if and only if $a - b = \zeta(c - b)$, where ζ is a primitive sixth root of unity.

Things To Watch Out For!

WARNING!!

The orientations of rotations matter in complex numbers, just as they matter in geometry.

Problem Solving Strategies

Concepts: 

- When placing a geometric problem on the complex plane, we can choose where to place the origin. Choosing a convenient point can often simplify the problem greatly.
- Sometimes we can prove a statement by proving a more general statement.
- Maintaining symmetry in systems of equations helps us find effective ways to work with the equations.
- A little wishful thinking can go a long way.
- Some complex number problems that don't appear to be about geometry can be solved quickly by thinking about the problem geometrically.
- When your initial approach to a problem hits a seemingly unsurpassable obstacle, take a step back and look for another way to view the problem that allows you to avoid the obstacle.
- Complex numbers can often be helpful in problems involving rotations, since rotations are easy to express with complex numbers.
- Working backwards from the desired result is a powerful problem solving strategy.
- Don't always specify the origin right away—sometimes waiting will allow you to exploit symmetry in a problem or will reveal a surprisingly useful choice for the origin.
- When applying complex numbers to a geometry problem involving a single circle, consider letting the circle be the unit circle centered at the origin.

REVIEW PROBLEMS

8.32 Describe the transformation that maps z to $\bar{z}$ in the complex plane.

8.33 For all z, let $f_1(z)$ be the image of z upon a translation of 2 units horizontally and -3 units vertically, and let $f_2(z)$ be the image of z upon a rotation about $4i$ counterclockwise by $\frac{\pi}{2}$.

(a) Find $(f_2 \circ f_1)(z)$ in terms of z.

(b) Find $(f_1 \circ f_2)(z)$ in terms of z.

8.34 Let $f(z) = i\bar{z}$. Describe the transformation that maps z to $f(z)$ for all z.

8.35 What must be true about the points corresponding to complex numbers a, b and c under each of the following conditions:

(a) $|a - b| + |a - c| = |b - c|$

(b) $|a - b|^2 + |a - c|^2 = |b - c|^2$

8.36 For a given nonzero complex number z, describe in words how to find the following points in the complex plane:

(a) $\dfrac{z}{|z|}$ (b) $z + \bar{z}$ (c) $z + |z|$

8.37 Use complex numbers to show that the line segment connecting the midpoints of two sides of a triangle is parallel to the third side and half the length of the third side.

8.38 Let A correspond to $3 - 2i$ and B correspond to $1 - 5i$.

(a) $\overline{AB}$ is perpendicular to the line through B and which of the following on the complex plane: $2 - 7i$, $-6 + 2i$, $-5 - i$?

(b) Find an equation whose graph consists of point B as well as all complex numbers z such that the line through z and B is perpendicular to $\overline{AB}$.

8.39 Use complex numbers to show that if $\angle ADB = 90°$, then point D is on the circle with $\overline{AB}$ as diameter.

8.40 Suppose that $2 + i$ and $8 - 3i$ are opposite vertices of a regular hexagon in the complex plane. What are the complex numbers that correspond to the other four vertices?

8.41 The median of a trapezoid is the segment with the midpoints of the trapezoid's legs as endpoints. Use complex numbers to prove that the median of a trapezoid is parallel to the bases of the trapezoid, and that the length of the median is the average of the lengths of the bases.

8.42 Use complex numbers to prove that the diagonals of a parallelogram are perpendicular if and only if the parallelogram is a rhombus.

8.43 Let a and b be distinct complex numbers. Show that z lies on the perpendicular bisector of a and b if and only if

$$(\bar{a} - \bar{b})z + (a - b)\bar{z} = a\bar{a} - b\bar{b}.$$

8.44 Suppose $WXYZ$ is a square with center O. Let T be the midpoint of $\overline{WX}$ and S be the midpoint of $\overline{OY}$. Use complex numbers to show that $\triangle TSZ$ is an isosceles right triangle.

8.45 Use complex numbers to show that if W, X, Y, and Z are four points in a plane such that $\overline{WX}$ is perpendicular to both $\overline{XY}$ and $\overline{WZ}$, then the midpoint of $\overline{YZ}$ is the same distance from W and X.

8.46 Show that a line segment joining the midpoints of opposite sides of a quadrilateral and the line segment joining the midpoints of the diagonals bisect each other.

8.47 Suppose $ABCD$ is a parallelogram and that complex numbers a, b, c, and d correspond to A, B, C, and D, respectively. Find d in terms of a, b, and c.

8.48 Show that $|w - z| \geq |w| - |z|$ for all complex numbers w and z.

8.49 Let P be a point inside rectangle $ABCD$. Show that $AP^2 + CP^2 = BP^2 + DP^2$.

Challenge Problems

8.50 Let O be the center of square $ABCD$. Show that no matter what line k we draw through O, the sum of the squares of the distances from the vertices of $ABCD$ to k is the same. **Hints:** 17

8.51 Given a triangle ABC, suppose we construct squares $BCDJ$, $CAFG$, and $ABHK$ on the sides of the triangle such that the squares do not overlap the triangle. We then locate points P and Q such that $GCDQ$ and $BHPJ$ are parallelograms. Prove that $\triangle PAQ$ is an isosceles right triangle. **Hints:** 269

8.52 Let z, a, and b be complex numbers such that $a \neq b$. Let w be the reflection of z over the line joining a and b, and let p be the projection of z onto the line joining a and b. (That is, let p be the foot of the perpendicular from z to the line through a and b.) Find p and w in terms of a, b, and z. **Hints:** 150

8.53 Let $f(x) = \sqrt{x^2 + (x-1)^2} + \sqrt{x^2 + (x+1)^2}$. Find the minimum value of f over all real numbers. **Hints:** 207, 40

8.54 If z_1 and z_2 are complex numbers such that $|z_1 - 5 + 3i| \leq 4$ and $|z_2 - 5i| \leq 2$, find the maximum and minimum values of $|z_1 - z_2|$. **Hints:** 35

8.55 Let z be a complex number. Find the minimum value of

$$|z-2| + |z-(1+2i)| + |z-(-1+i)| + |z-(-1-i)|.$$

Hints: 151

8.56 Let $\triangle ABC$ be a triangle in the complex plane whose vertices correspond to complex numbers a, b, and c. Show that each of the following is a necessary and sufficient condition for $\triangle ABC$ to be equilateral:

(a) $a^2 + b^2 + c^2 = ab + ac + bc$. **Hints:** 111, 233

(b) $\frac{1}{a-b} + \frac{1}{b-c} + \frac{1}{c-a} = 0$.

(c) $(b-c)^2 + (c-a)^2 + (a-b)^2 = 0$.

8.57★ For a positive integer n, define S_n to be the minimum value of the sum

$$\sqrt{(2\cdot 1-1)^2 + a_1^2} + \sqrt{(2\cdot 2-1)^2 + a_2^2} + \sqrt{(2\cdot 3-1)^2 + a_3^2} + \cdots + \sqrt{(2\cdot n-1)^2 + a_n^2},$$

where $a_1, a_2, \ldots, a_n$ are positive real numbers whose sum is 17. There is a unique positive integer n for which S_n is also an integer. Find this n. *(Source: AIME)*

8.58 Suppose $ABCDE$ is a pentagon such that the lines connecting A, B, C, and D respectively to the midpoints of $\overline{EB}$, $\overline{AC}$, $\overline{BD}$, and $\overline{CE}$ are concurrent. Show that the line connecting E to the midpoint of $\overline{AD}$ also passes through this common point. **Hints:** 95, 264

8.59 Let $P_0P_1\cdots P_{n-1}$ be a regular n-gon inscribed in a unit circle C, and let P be a point on C.

(a) Prove that $PP_0^2 + PP_1^2 + \cdots + PP_{n-1}^2 = 2n$ for any P. **Hints:** 172

(b) Prove that $PP_0^4 + PP_1^4 + \cdots + PP_{n-1}^4 = 6n$ for any P.

8.60★ There are $2n$ complex numbers that satisfy both $z^{28} - z^8 - 1 = 0$ and $|z| = 1$. These numbers have the form $z_m = \cos\theta_m + i\sin\theta_m$, where $0° \le \theta_1 < \theta_2 < \cdots < \theta_{2n} < 360°$ and angles are measured in degrees. Find the value of $\theta_2 + \theta_4 + \cdots + \theta_{2n}$. *(Source: AIME)* **Hints:** 22, 113

8.61 Regular decagon $P_1P_2\cdots P_{10}$ is drawn in the coordinate plane, with P_1 at $(1,0)$ and P_6 at $(3,0)$. If P_n is the point (x_n, y_n), compute the numerical value of

$$(x_1 + y_1i)(x_2 + y_2i)(x_3 + y_3i)\cdots(x_{10} + y_{10}i).$$

(Source: ARML) **Hints:** 217

8.62★ The sides of rectangle $ABCD$ have lengths 10 and 11. An equilateral triangle is drawn so that no point of the triangle lies outside $ABCD$. Find the maximum possible area of such a triangle. *(Source: AIME)* **Hints:** 161

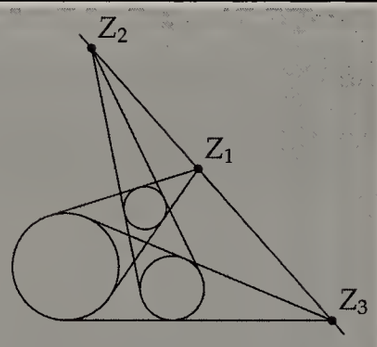
8.63★ Let P be a point inside equilateral triangle ABC with side s. Show that

$$3(PA^4 + PB^4 + PC^4 + s^4) = (PA^2 + PB^2 + PC^2 + s^2)^2.$$

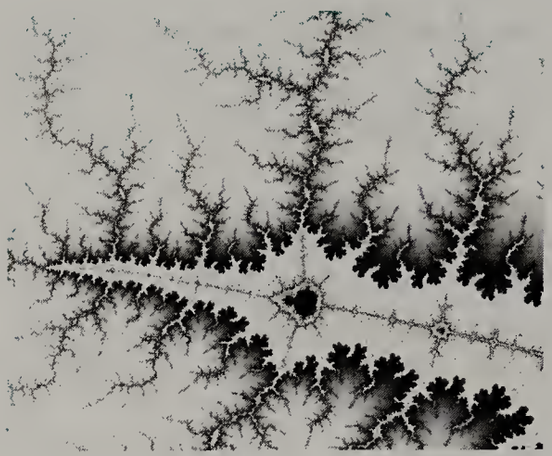
Hints: 96, 89

8.64★ Given triangle ABC, construct equilateral triangles BCX, CAY, ABZ externally. Let B' and N' be the midpoints of $\overline{BC}$ and $\overline{CY}$, respectively, and let M and N be the midpoints of $\overline{AZ}$ and $\overline{CX}$, respectively. Prove that $\overline{B'N'} \perp \overline{MN}$. *(Source: Argentina)* **Hints:** 127, 85

Extra! Here's the magical three-dimensional proof of Monge's Theorem that is suggested on page 289. Suppose we start with three spheres with different radii such that no sphere is entirely within another. There are two planes that are tangent to all three spheres such that the three spheres are sandwiched between the two planes. Since the spheres do not have the same radii, these two planes intersect in a line which we'll call line k. For each pair of spheres, there is a cone that is tangent to both spheres. The aforementioned planes are also tangent to each cone, and therefore include the vertex of each cone. So, the vertex of each of the three cones determined by a pair of the spheres is on the line that is the intersection of the planes.



With all that mind, picture the cross-section that is formed by the plane that bisects the angle between the planes. This plane must pass through the centers of the spheres, and must include line k (and therefore the vertex of each cone). Moreover, the intersection of the plane and each cone produces the two lines that are tangent to the corresponding pair of circles. Therefore, this cross-section produces exactly the situation in Monge's Theorem, and shows why the three points in question are collinear.



A plan is just a tangent vector on the manifold of reality. – Skratch Garrison

CHAPTER 9
Vectors in Two Dimensions

9.1 What's a Vector?

A **vector in two dimensions** is an ordered pair of numbers, which we often denote as $\begin{pmatrix} x \\ y \end{pmatrix}$. The numbers x and y are called **components** of the vector. Vectors have the following operations:

- **Vector addition**. We add vectors by adding the corresponding components:

$$\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a+c \\ b+d \end{pmatrix}.$$

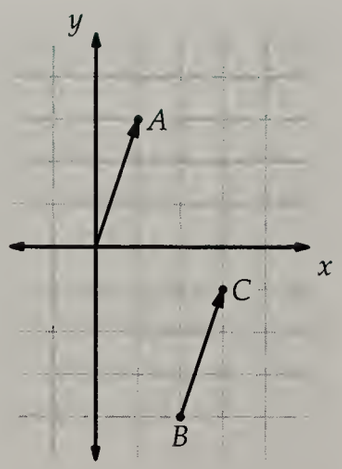
- **Multiplication of a vector by a constant**. We multiply a vector by a constant by multiplying each component of the vector by the constant:

$$k\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ka \\ kb \end{pmatrix}.$$

We often refer to real numbers as **scalars** when speaking of vectors. So, the product $k\begin{pmatrix} a \\ b \end{pmatrix}$ is commonly referred to as **scalar multiplication**. In Problem 9.4, we'll see why real numbers earn this name when speaking of vectors.

Instead of referring to a vector by its components, we can give a vector a label. When we do so, we usually use a bold lowercase letter that is not italicized, like $\mathbf{v}$. Other notations that are sometimes used to denote a vector are $\vec{v}$ or $\hat{v}$.

A vector can be represented geometrically as an arrow. The arrow from $(0,0)$ to $(1,3)$ in the diagram at right represents the vector $\begin{pmatrix}1\\3\end{pmatrix}$. The diagram also shows an arrow from $(2,-4)$ to $(3,-1)$. *These arrows represent the same vector.* The location of a vector is irrelevant. So, any arrow that depicts "right 1 unit, up 3 units" represents the vector $\begin{pmatrix}1\\3\end{pmatrix}$.



We will occasionally refer to vectors with points, writing $\vec{BC}$ for the vector that is represented by the arrow from B to C. If we write a vector with one point, such as $\vec{A}$, we mean the vector from the origin to that point. (Note that the notation $\vec{BC}$ is slightly different than the notation $\overrightarrow{BC}$. $\overrightarrow{BC}$ denotes the ray that starts at B, passes through C, *and keeps going forever in that direction.* The vector $\vec{BC}$ is represented by an arrow that goes from B to C and stops there.)

We won't always be this precise in our language when referring to vectors in a geometric context. We will often simply refer to "an arrow that represents a vector" as simply "a vector," rather than littering the text with "an arrow that represents." So, letting $\mathbf{v} = \begin{pmatrix}1\\3\end{pmatrix}$, we see that both $\vec{A}$ and $\vec{BC}$ in the diagram above are $\mathbf{v}$, and we can write $\vec{A} = \vec{BC} = \mathbf{v}$.

A vector with magnitude 1 is called a **unit vector**. The vector $\begin{pmatrix}x\\y\end{pmatrix}$ can also be written as $x\mathbf{i} + y\mathbf{j}$, where $\mathbf{i}$ is the unit vector $\begin{pmatrix}1\\0\end{pmatrix}$ and $\mathbf{j}$ is the unit vector $\begin{pmatrix}0\\1\end{pmatrix}$. For example, the vectors $\vec{A}$ and $\vec{BC}$ shown at right above are both $\mathbf{i} + 3\mathbf{j}$. We will mostly use the notation $\begin{pmatrix}1\\3\end{pmatrix}$, but will occasionally mix in $\mathbf{i} + 3\mathbf{j}$ notation so you'll be comfortable with it.

The **zero vector** is the vector in which both components are zero. We often denote this vector as simply $\mathbf{0}$. We also have a special symbol for the set of all vectors in two dimensions: $\mathbb{R}^2$. To see why we use this symbol, recall that $\mathbb{R}$ is the set of all real numbers. So, we use the symbol $\mathbb{R}^2$ to indicate the set of all ordered pairs of real numbers.

Problems

Problem 9.1: We define the **norm** of a vector $\begin{pmatrix}x\\y\end{pmatrix}$ to be the distance from (x,y) to the origin in the Cartesian plane. We denote the norm of a vector $\mathbf{v}$ as $\|\mathbf{v}\|$. Find $\left\|\begin{pmatrix}x\\y\end{pmatrix}\right\|$ in terms of x and y.

Problem 9.2:

(a) If $\mathbf{v} = 3\mathbf{i} - 4\mathbf{j}$, then what is $\left\|\dfrac{\mathbf{v}}{\|\mathbf{v}\|}\right\|$?

(b) If $\mathbf{v} = \begin{pmatrix}7\\-7\end{pmatrix}$, then what is $\left\|\dfrac{\mathbf{v}}{\|\mathbf{v}\|}\right\|$?

(c) If $\mathbf{v} = \begin{pmatrix}0.5\\-0.2\end{pmatrix}$, then what is $\left\|\dfrac{\mathbf{v}}{\|\mathbf{v}\|}\right\|$?

(d) Generalize your results from the first three parts. That is, what can we say about $\left\|\dfrac{\mathbf{v}}{\|\mathbf{v}\|}\right\|$ when $\mathbf{v}$ is a nonzero vector?

Problem 9.3: Let $x\mathbf{i} + y\mathbf{j}$ be a nonzero vector, and let (r, θ), where $r > 0$, be the polar coordinates of the point with rectangular coordinates (x, y). How are r and θ related to the vector $x\mathbf{i} + y\mathbf{j}$?

Problem 9.4: Let $\mathbf{v}$ be a nonzero vector.

(a) Geometrically speaking, how are $\mathbf{v}$ and $3\mathbf{v}$ related?

(b) Let $-\mathbf{v} = (-1)\mathbf{v}$ for any $\mathbf{v}$. Geometrically speaking, how are $\mathbf{v}$ and $-\mathbf{v}$ related?

Problem 9.5: In this problem, we explore vector addition. Let $\mathbf{v} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$.

(a) Find x and y such that $\mathbf{v} + \mathbf{w} = \begin{pmatrix} x \\ y \end{pmatrix}$.

(b) Let O be the origin, let A be the point $(5, 1)$, let B be the point $(2, 4)$, and let C be the point (x, y), where x and y are your answers from part (a). What are $\vec{OA}$, $\vec{OB}$, $\vec{OC}$, and $\vec{AC}$ in terms of $\mathbf{v}$ and $\mathbf{w}$? What type of quadrilateral is $OACB$?

(c) Use your answers to part (b) to give a geometric description of adding two vectors.

Problem 9.6: In this problem, we explore vector subtraction. Let $\mathbf{v} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$.

(a) We define vector subtraction to fit with our definition of vector addition. The vector $\mathbf{w} - \mathbf{v}$ is the vector $\mathbf{u}$ such that $\mathbf{u} + \mathbf{v} = \mathbf{w}$. Use the definition of vector addition to find x and y such that $\mathbf{w} - \mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$.

(b) Let O be the origin, let A be the point $(5, 1)$, let B be the point $(2, 4)$, and let C be the point (x, y), where x and y are your answers from part (a). What are $\vec{OA}$, $\vec{OB}$, $\vec{OC}$, and $\vec{AB}$ in terms of $\mathbf{v}$ and $\mathbf{w}$?

(c) Show that for any points P and Q in the Cartesian plane, we have $\vec{PQ} = \vec{Q} - \vec{P}$.

(d) How are $\vec{PQ}$ and $\vec{QP}$ related?

Problem 9.1: We define the **norm** of a vector $\begin{pmatrix} x \\ y \end{pmatrix}$ to be the distance from (x, y) to the origin in the Cartesian plane. We denote the norm of a vector $\mathbf{v}$ as $\|\mathbf{v}\|$. Find $\left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|$ in terms of x and y.

Solution for Problem 9.1: Applying the distance formula gives us

$$\left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\| = \sqrt{(x-0)^2 + (y-0)^2} = \sqrt{x^2 + y^2}.$$

□

Important: The **norm** of a vector $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ is

$$\|\mathbf{v}\| = \sqrt{x^2 + y^2}.$$

The norm of a vector is also called the **magnitude** or **length** of the vector, since it measures the length of the geometric representation of the vector.

Problem 9.2:

(a) If $\mathbf{v} = 3\mathbf{i} - 4\mathbf{j}$, then what is $\left\| \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\|$?

(b) If $\mathbf{v} = \begin{pmatrix} 7 \\ -7 \end{pmatrix}$, then what is $\left\| \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\|$?

(c) If $\mathbf{v} = \begin{pmatrix} 0.5 \\ -0.2 \end{pmatrix}$, then what is $\left\| \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\|$?

(d) Generalize your results from the first three parts. That is, what can we say about $\left\| \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\|$ when $\mathbf{v}$ is a nonzero vector?

Solution for Problem 9.2:

(a) We have $\|\mathbf{v}\| = \sqrt{3^2 + (-4)^2} = 5$, so $\frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{1}{5}\mathbf{v} = \frac{3}{5}\mathbf{i} - \frac{4}{5}\mathbf{j}$. Therefore, we have

$$\left\| \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\| = \sqrt{\left(\frac{3}{5}\right)^2 + \left(-\frac{4}{5}\right)^2} = 1.$$

(b) We have $\|\mathbf{v}\| = \sqrt{7^2 + (-7)^2} = 7\sqrt{2}$, so $\frac{\mathbf{v}}{\|\mathbf{v}\|} = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$ and

$$\left\| \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\| = \sqrt{\frac{1}{2} + \frac{1}{2}} = 1.$$

Very suspicious.

(c) We have $\|\mathbf{v}\| = \sqrt{(0.5)^2 + (-0.2)^2} = \sqrt{0.29}$, so $\frac{\mathbf{v}}{\|\mathbf{v}\|} = \begin{pmatrix} 0.5/\sqrt{0.29} \\ -0.2/\sqrt{0.29} \end{pmatrix}$. Rather than rationalizing the denominator or getting rid of the decimals, we forge ahead finding $\left\| \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\|$, because we're pretty confident something convenient will happen:

$$\left\| \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\| = \sqrt{\left(\frac{0.5}{\sqrt{0.29}}\right)^2 + \left(-\frac{0.2}{\sqrt{0.29}}\right)^2} = \sqrt{\frac{0.25 + 0.04}{0.29}} = 1.$$

Very convenient, indeed.

(d) We have a pretty good idea what the result will be. We let $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ and work through the details. We have $\|\mathbf{v}\| = \sqrt{x^2 + y^2}$, which is nonzero because $\mathbf{v} \neq \mathbf{0}$, so

$$\left\| \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\| = \sqrt{\left(\frac{x}{\sqrt{x^2 + y^2}}\right)^2 + \left(\frac{y}{\sqrt{x^2 + y^2}}\right)^2} = \sqrt{\frac{x^2 + y^2}{x^2 + y^2}} = 1.$$

□

As noted in the introduction, a vector with magnitude 1 is called a unit vector. In Problem 9.2, we showed that the result of dividing a nonzero vector by its magnitude is a unit vector. Since $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ is a unit vector, when we write $\mathbf{v}$ as

$$\mathbf{v} = \|\mathbf{v}\| \cdot \frac{\mathbf{v}}{\|\mathbf{v}\|},$$

we express $\mathbf{v}$ as the product of its magnitude and a unit vector. For example, we have $\|3\mathbf{i} - 4\mathbf{j}\| = \sqrt{3^2 + (-4)^2} = 5$, so we can express $3\mathbf{i} - 4\mathbf{j}$ as the product of its magnitude and direction vector as follows:

$$3\mathbf{i} - 4\mathbf{j} = 5\left(\frac{3}{5}\mathbf{i} - \frac{4}{5}\mathbf{j}\right).$$

Geometrically speaking, we can think of this unit vector as providing the direction of the vector, and the magnitude as telling us how far the vector goes in that direction. Since each nonzero vector can be written as the product of its magnitude and a unit vector, we can think of a nonzero vector as being defined by its magnitude and its direction (as given by the unit vector).

We can also define the direction of a vector by using an angle, as we'll see in the next problem.

Problem 9.3: Let $x\mathbf{i} + y\mathbf{j}$ be a nonzero vector, and let (r, θ), where $r > 0$, be the polar coordinates of the point with rectangular coordinates (x, y). How are r and θ related to the vector $x\mathbf{i} + y\mathbf{j}$?

Solution for Problem 9.3: Since $x = r\cos\theta$ and $y = r\sin\theta$, we have

$$x\mathbf{i} + y\mathbf{j} = (r\cos\theta)\mathbf{i} + (r\sin\theta)\mathbf{j} = r((\cos\theta)\mathbf{i} + (\sin\theta)\mathbf{j}).$$

We have

$$r = \sqrt{x^2 + y^2} = \|x\mathbf{i} + y\mathbf{j}\|,$$

so r is the magnitude of $x\mathbf{i} + y\mathbf{j}$. Since

$$\|(\cos\theta)\mathbf{i} + (\sin\theta)\mathbf{j}\| = \sqrt{(\cos\theta)^2 + (\sin\theta)^2} = 1,$$

we see that $(\cos\theta)\mathbf{i} + (\sin\theta)\mathbf{j}$ is a unit vector, so θ can be used to describe the direction of the vector. □

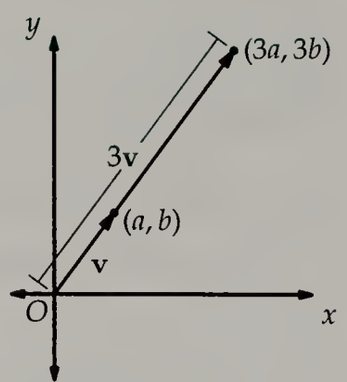
Problem 9.4: Let $\mathbf{v}$ be a nonzero vector.

(a) Geometrically speaking, how are $\mathbf{v}$ and $3\mathbf{v}$ related?

(b) Let $-\mathbf{v} = (-1)\mathbf{v}$ for any $\mathbf{v}$. Geometrically speaking, how are $\mathbf{v}$ and $-\mathbf{v}$ related?

Solution for Problem 9.4:

(a) If we let $\mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix}$, we have $3\mathbf{v} = \begin{pmatrix} 3a \\ 3b \end{pmatrix}$. The origin, the point (a, b), and the point $(3a, 3b)$ are collinear, since all lie on the line through the origin with slope $\frac{b}{a}$ (when a is nonzero; if $a = 0$, then all three points are on the y-axis). Therefore, it appears that $\mathbf{v}$ and $3\mathbf{v}$ have the same direction. To be more precise, we can express each as the product of its magnitude and direction. We have $\|\mathbf{v}\| = \sqrt{a^2 + b^2}$ and $\|3\mathbf{v}\| = \sqrt{(3a)^2 + (3b)^2} = 3\sqrt{a^2 + b^2}$, so



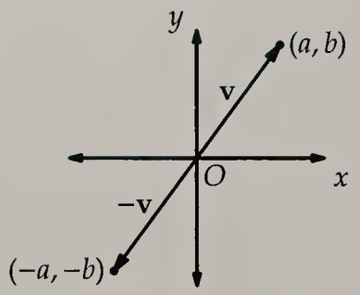
$$\mathbf{v} = \sqrt{a^2 + b^2}\left(\frac{a}{\sqrt{a^2 + b^2}}\mathbf{i} + \frac{b}{\sqrt{a^2 + b^2}}\mathbf{j}\right),$$

$$3\mathbf{v} = 3\sqrt{a^2 + b^2}\left(\frac{3a}{3\sqrt{a^2 + b^2}}\mathbf{i} + \frac{3b}{3\sqrt{a^2 + b^2}}\mathbf{j}\right)$$

$$= 3\sqrt{a^2 + b^2}\left(\frac{a}{\sqrt{a^2 + b^2}}\mathbf{i} + \frac{b}{\sqrt{a^2 + b^2}}\mathbf{j}\right).$$

We see that the vector $3\mathbf{v}$ has same direction as $\mathbf{v}$, but has three times the magnitude. That is, $3\mathbf{v}$ is a "scaled" version of $\mathbf{v}$.

(b) If $\mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix}$, then $-\mathbf{v} = \begin{pmatrix} -a \\ -b \end{pmatrix}$. These two vectors are indicated in the diagram at right. The point (a, b) and the point $(-a, -b)$ are on the line through the origin with slope $\frac{b}{a}$ (when $a \neq 0$; if $a = 0$, then all three points on the y-axis with the origin in the middle). However, the origin is between (a, b) and $(-a, -b)$, so the vector from the origin to (a, b) is in the opposite direction as the vector from the origin to $(-a, -b)$. Since the origin is equidistant from (a, b) and $(-a, -b)$, we see that $\mathbf{v}$ and $-\mathbf{v}$ have the same magnitude, but opposite direction.



□

Problem 9.5: In this problem, we explore vector addition. Let $\mathbf{v} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$.

(a) Find x and y such that $\mathbf{v} + \mathbf{w} = \begin{pmatrix} x \\ y \end{pmatrix}$.

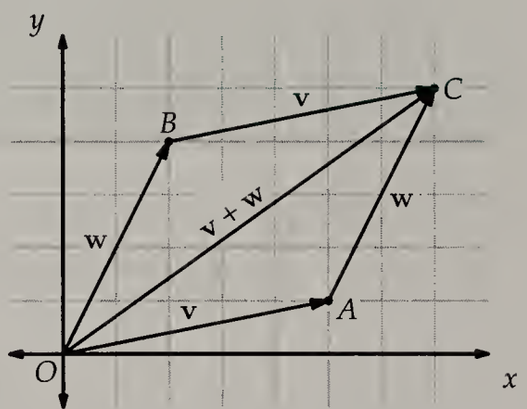
(b) Let O be the origin, let A be the point $(5, 1)$, let B be the point $(2, 4)$, and let C be the point (x, y), where x and y are your answers from part (a). What are $\vec{OA}$, $\vec{OB}$, $\vec{OC}$, and $\vec{AC}$ in terms of $\mathbf{v}$ and $\mathbf{w}$? What type of quadrilateral is $OACB$?

(c) Use your answers to part (b) to give a geometric description for adding two vectors.

Solution for Problem 9.5:

(a) Following the rule for vector addition, we have $\begin{pmatrix} 5 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 5+2 \\ 1+4 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$.

(b) Since the arrow from O to A indicates going to the right 5 and up 1, we have $\vec{OA} = \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \mathbf{v}$. Similarly, we have $\vec{OB} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \mathbf{w}$, and $\vec{OC} = \begin{pmatrix} 7 \\ 5 \end{pmatrix} = \mathbf{v} + \mathbf{w}$. To get from A to C, we go right 2 and up 4, so $\vec{AC} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$. This is $\mathbf{w}$! We therefore have $\vec{AC} = \mathbf{w} = \vec{OB}$, which means sides $\overline{AC}$ and $\overline{OB}$ of $OACB$ are parallel ($\vec{AC}$ and $\vec{OB}$ have the same direction) and have the same length ($\vec{AC}$ and $\vec{OB}$ have the same magnitude). So, $OACB$ is a parallelogram.



(c) In $\triangle OAC$, we see that $\vec{OA} = \mathbf{v}$, $\vec{AC} = \mathbf{w}$, and $\vec{OC} = \mathbf{v} + \mathbf{w}$. This isn't a coincidence. We can follow exactly the same steps to add any two vectors. We start at the origin and draw the arrow corresponding to $\mathbf{v}$. Then, starting from where this arrow ends, we draw the arrow corresponding to $\mathbf{w}$. The arrow from the origin to the end of this second arrow corresponds to $\mathbf{v} + \mathbf{w}$. We can also view $\mathbf{v} + \mathbf{w}$ as the diagonal of the parallelogram with $\mathbf{v}$ and $\mathbf{w}$ as consecutive sides.

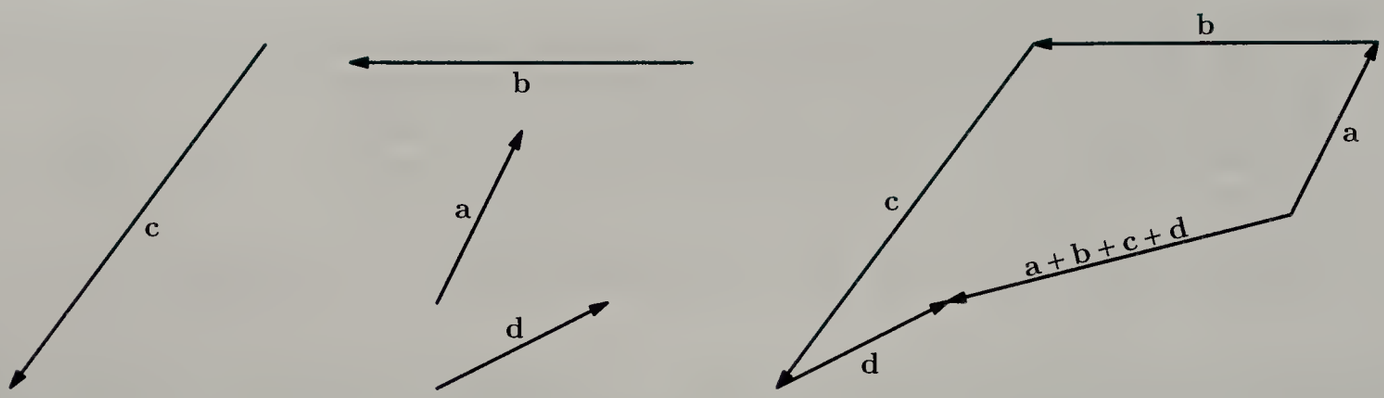
We can also see in our diagram that vector addition is commutative. That is, we have

$$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v},$$

where $\mathbf{v} + \mathbf{w}$ corresponds to the path $O \to A \to C$ and $\mathbf{w} + \mathbf{v}$ corresponds to the path $O \to B \to C$.

□

The vector addition illustrated in the previous problem can be extended to any number of vectors. The process of adding vectors described in part (c) is sometimes referred to as **head-to-tail addition**, where the start of an arrow depicting a vector is the **tail** of the vector and the end of the arrow is the vector's **head**.



We add a series of vectors by drawing the first, and then drawing the second by placing the tail of the second at the head of the first. We continue in this manner until all the vectors are drawn. The arrow from the tail of the first vector to the head of the last represents the sum of all the vectors.

Now that we have a handle on addition, we turn to vector subtraction.

Problem 9.6: Let $\mathbf{v} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$.

(a) We define vector subtraction to fit with our definition of vector addition. The vector $\mathbf{w} - \mathbf{v}$ is the vector $\mathbf{u}$ such that $\mathbf{u} + \mathbf{v} = \mathbf{w}$. Use the definition of vector addition to find x and y such that $\mathbf{w} - \mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$.

(b) Let O be the origin, let A be the point $(5, 1)$, let B be the point $(2, 4)$, and let C be the point (x, y), where x and y are your answers from part (a). What are $\vec{OA}$, $\vec{OB}$, $\vec{OC}$, and $\vec{AB}$ in terms of $\mathbf{v}$ and $\mathbf{w}$?

(c) Show that for any points P and Q in the Cartesian plane, we have $\vec{PQ} = \vec{Q} - \vec{P}$.

(d) How are $\vec{PQ}$ and $\vec{QP}$ related?

Solution for Problem 9.6:

(a) If $\mathbf{w} - \mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$, then we must have $\mathbf{w} = \begin{pmatrix} x \\ y \end{pmatrix} + \mathbf{v}$. Applying the definition of vector addition, we have

$$\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} x+5 \\ y+1 \end{pmatrix}.$$

Therefore, we must have $x + 5 = 2$ and $y + 1 = 4$, so $x = 2 - 5 = -3$ and $y = 4 - 1 = 3$. In other words, vector subtraction works in exactly the way we would expect:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{w} - \mathbf{v} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} - \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 2-5 \\ 4-1 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \end{pmatrix}.$$

In general, we can follow the same process to show that:

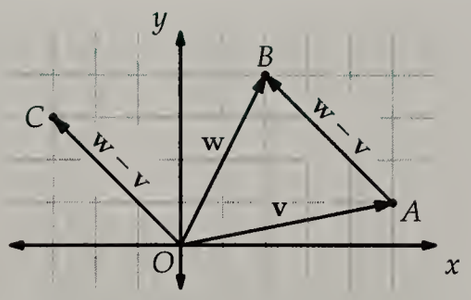
Important: For any vectors $\begin{pmatrix} a \\ b \end{pmatrix}$ and $\begin{pmatrix} c \\ d \end{pmatrix}$, we have

$$\begin{pmatrix} a \\ b \end{pmatrix} - \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a-c \\ b-d \end{pmatrix}.$$

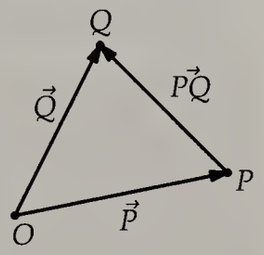
This also means that we can view vector subtraction as a combination of vector addition and scalar multiplication:

$$\begin{pmatrix} a \\ b \end{pmatrix} - \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} + (-1)\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} -c \\ -d \end{pmatrix} = \begin{pmatrix} a-c \\ b-d \end{pmatrix}.$$

(b) As in the previous problem, we have $\vec{OA} = \mathbf{v}$ and $\vec{OB} = \mathbf{w}$. Since $(x, y) = (-3, 3)$ from part (a), we have $\vec{OC} = \begin{pmatrix} -3 \\ 3 \end{pmatrix} = \mathbf{w} - \mathbf{v}$. Since point B is 3 to the left and 3 above A, we have $\vec{AB} = \begin{pmatrix} -3 \\ 3 \end{pmatrix}$, so $\vec{AB} = \mathbf{w} - \mathbf{v}$.



(c) In the diagram, let O be the origin. From our definition of vector subtraction, the vector $\vec{Q} - \vec{P}$ is the vector we must add to $\vec{P}$ to get $\vec{Q}$. Consider the vector sum $\vec{P} + \vec{PQ}$. Since $\vec{P}$ is the vector from the origin to P, and the vector $\vec{PQ}$ is the vector from P to Q, when we perform head-to-tail addition to get $\vec{P} + \vec{PQ}$, we go from the origin to P, and then from P to Q. So, the vector sum $\vec{P} + \vec{PQ}$ equals $\vec{Q}$. Therefore, $\vec{PQ}$ is indeed the vector we must add to $\vec{P}$ to get $\vec{Q}$, which means $\vec{PQ} = \vec{Q} - \vec{P}$.



(d) Using the relationship from part (c), we have $\vec{PQ} = \vec{Q} - \vec{P}$ and $\vec{QP} = \vec{P} - \vec{Q}$, so $\vec{PQ} = -\vec{QP}$. We could also have noted that the vector from P to Q (the vector $\vec{PQ}$) has the same length but opposite direction as the vector from Q to P ($\vec{QP}$), so we have $\vec{PQ} = -\vec{QP}$.

□

Important: For any points A and B, we have $\vec{AB} = \vec{B} - \vec{A}$.

Exercises

9.1.1 Use the definitions of vector addition and multiplication of a vector by a scalar to show that $\mathbf{0} + \mathbf{v} = \mathbf{v}$ and $0\mathbf{v} = \mathbf{0}$ for all $\mathbf{v}$.

9.1.2 Evaluate the following:

(a) $\left\| \begin{pmatrix} 4 \\ -5 \end{pmatrix} \right\|$ (b) $\|7\mathbf{i} - 24\mathbf{j}\|$

9.1.3 Express each of the following as a single vector:

(a) $\begin{pmatrix} 4 \\ -3 \end{pmatrix} + 2\begin{pmatrix} 5 \\ -3 \end{pmatrix} - \begin{pmatrix} -1 \\ 7 \end{pmatrix}$ (b) $\begin{pmatrix} -1 \\ 6 \end{pmatrix} - 2\begin{pmatrix} -6 \\ 2 \end{pmatrix} - 3\begin{pmatrix} 0 \\ 4 \end{pmatrix}$

9.1.4 If A, B, C, and D are points in the Cartesian plane such that $\vec{AB} = \vec{CD}$ and the four points are not all on the same line, then must $ABDC$ be a parallelogram?

9.1.5 Show that $\|c\mathbf{v}\| = |c|\,\|\mathbf{v}\|$ for all vectors $\mathbf{v}$ and scalars c.

9.2 The Dot Product

In the previous section, we learned how to add and subtract vectors and how to multiply a vector by a scalar. In this section, we define a special product of two vectors, the **dot product**. We denote the dot product of vectors $\mathbf{v}$

and $\mathbf{w}$ as $\mathbf{v} \cdot \mathbf{w}$. Letting $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$, we define the dot product $\mathbf{v} \cdot \mathbf{w}$ as

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = v_1 w_1 + v_2 w_2.$$

The dot product is also sometimes referred to as the **inner product**.

Problems

Problem 9.7: Let $\mathbf{p} = \begin{pmatrix} 4 \\ -3 \end{pmatrix}$, $\mathbf{q} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$, and $\mathbf{r} = \begin{pmatrix} -2 \\ 7 \end{pmatrix}$.

(a) Evaluate $\mathbf{p} \cdot \mathbf{q}$ and $\mathbf{q} \cdot \mathbf{p}$.

(b) Evaluate $\mathbf{p} \cdot (2\mathbf{r})$ and $2(\mathbf{p} \cdot \mathbf{r})$.

(c) Evaluate $\mathbf{p} \cdot (\mathbf{q} + \mathbf{r})$ and $\mathbf{p} \cdot \mathbf{q} + \mathbf{p} \cdot \mathbf{r}$.

(d) Generalize your results from parts (a), (b), and (c). That is, show that for all vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ and all scalars a, we have $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$, $\mathbf{u} \cdot (a\mathbf{v}) = a(\mathbf{u} \cdot \mathbf{v})$ and $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$.

Problem 9.8: Show that for any vector $\mathbf{v}$, we have $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$.

Problem 9.9: In this problem, we find a geometric interpretation of the dot product of two nonzero vectors.

(a) Let V and W be points in the Cartesian plane such that $\vec{V} = \mathbf{v}$ and $\vec{W} = \mathbf{w}$, and let O be the origin. Let $\theta = \angle WOV$. Use $\triangle WOV$ to express $(\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w})$ in terms of $\|\mathbf{v}\|$, $\|\mathbf{w}\|$, and θ.

(b) Expand the product $(\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w})$ using your results from Problem 9.7.

(c) Combine your results in parts (a) and (b) to show that $\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \, \|\mathbf{w}\| \cos\theta$.

Problem 9.10: Find the angle between each pair of vectors below.

(a) $\begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix}$ and $\begin{pmatrix} 0 \\ -2 \end{pmatrix}$

(b) $\begin{pmatrix} 12 \\ -8 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$

(c) $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -2 - \sqrt{3} \\ -1 + 2\sqrt{3} \end{pmatrix}$

Problem 9.11: We say that two vectors are **orthogonal** if their dot product is zero. Suppose we draw arrows representing two nonzero vectors from the same point. If these vectors are orthogonal, then how are the arrows that represent them related?

Problem 9.12: If $\mathbf{p} \cdot \mathbf{q} = \mathbf{p} \cdot \mathbf{r}$ for a nonzero vector $\mathbf{p}$, then must we have $\mathbf{q} = \mathbf{r}$?

Problem 9.13: Use the dot product to prove that

$$(x_1 y_1 + x_2 y_2)^2 \leq (x_1^2 + x_2^2)(y_1^2 + y_2^2)$$

for all real x_1, x_2, y_1, and y_2. When does equality hold?

Problem 9.14: Suppose f is a function whose domain is $\mathbb{R}^2$, which means that its domain is all vectors in two dimensions. We say that f is a **linear function** if

(i) $f(a\mathbf{v}) = af(\mathbf{v})$ for all scalars a and vectors $\mathbf{v}$, and

(ii) $f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w})$ for all vectors $\mathbf{v}$ and $\mathbf{w}$.

Suppose $g(\mathbf{w})$ is a linear function that maps vectors to scalars. That is, it takes a vector as an input and gives a scalar as an output. For example, $g(\mathbf{w}) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \cdot \mathbf{w}$ is one such function.

(a) Explain why $g(x\mathbf{i} + y\mathbf{j}) = xg(\mathbf{i}) + yg(\mathbf{j})$.

(b) Why does part (a) tell us that there exists some vector $\mathbf{v}$ such that $g(\mathbf{w}) = \mathbf{v} \cdot \mathbf{w}$ for all $\mathbf{w}$? Is this vector $\mathbf{v}$ unique?

Problem 9.7: Let $\mathbf{p} = \begin{pmatrix} 4 \\ -3 \end{pmatrix}$, $\mathbf{q} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$, and $\mathbf{r} = \begin{pmatrix} -2 \\ 7 \end{pmatrix}$.

(a) Evaluate $\mathbf{p} \cdot \mathbf{q}$ and $\mathbf{q} \cdot \mathbf{p}$.

(b) Evaluate $\mathbf{p} \cdot (2\mathbf{r})$ and $2(\mathbf{p} \cdot \mathbf{r})$.

(c) Evaluate $\mathbf{p} \cdot (\mathbf{q} + \mathbf{r})$ and $\mathbf{p} \cdot \mathbf{q} + \mathbf{p} \cdot \mathbf{r}$.

(d) Generalize your results from parts (a), (b), and (c). That is, show that for all vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ and all scalars a, we have $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$, $\mathbf{u} \cdot (a\mathbf{v}) = a(\mathbf{u} \cdot \mathbf{v})$ and $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$.

Solution for Problem 9.7:

(a) We have $\mathbf{p} \cdot \mathbf{q} = (4)(5) + (-3)(2) = 14$ and $\mathbf{q} \cdot \mathbf{p} = (5)(4) + (2)(-3) = 14$, so $\mathbf{p} \cdot \mathbf{q} = \mathbf{q} \cdot \mathbf{p}$.

(b) We have

$$\mathbf{p} \cdot (2\mathbf{r}) = \begin{pmatrix} 4 \\ -3 \end{pmatrix} \cdot \begin{pmatrix} -4 \\ 14 \end{pmatrix} = (4)(-4) + (-3)(14) = -58$$

and

$$2(\mathbf{p} \cdot \mathbf{r}) = 2\left(\begin{pmatrix} 4 \\ -3 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ 7 \end{pmatrix}\right) = 2(4(-2) + (-3)(7)) = -58.$$

Interesting; we have $\mathbf{p} \cdot (2\mathbf{r}) = 2(\mathbf{p} \cdot \mathbf{r})$.

(c) We have

$$\mathbf{p} \cdot (\mathbf{q} + \mathbf{r}) = \begin{pmatrix} 4 \\ -3 \end{pmatrix} \cdot \left(\begin{pmatrix} 5 \\ 2 \end{pmatrix} + \begin{pmatrix} -2 \\ 7 \end{pmatrix}\right) = \begin{pmatrix} 4 \\ -3 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 9 \end{pmatrix} = (4)(3) + (-3)(9) = -15$$

and

$$\mathbf{p} \cdot \mathbf{q} + \mathbf{p} \cdot \mathbf{r} = \begin{pmatrix} 4 \\ -3 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ 2 \end{pmatrix} + \begin{pmatrix} 4 \\ -3 \end{pmatrix} \begin{pmatrix} -2 \\ 7 \end{pmatrix} = 14 - 29 = -15.$$

We're not surprised to see that $\mathbf{p} \cdot (\mathbf{q} + \mathbf{r}) = \mathbf{p} \cdot \mathbf{q} + \mathbf{p} \cdot \mathbf{r}$.

(d) Let $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, and $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$. We then have

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2,$$
$$\mathbf{v} \cdot \mathbf{u} = v_1u_1 + v_2u_2 = u_1v_1 + u_2v_2,$$

so the dot product is commutative: we have $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ for all vectors $\mathbf{u}$ and $\mathbf{v}$. We also have

$$\mathbf{u} \cdot (a\mathbf{v}) = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \cdot \begin{pmatrix} av_1 \\ av_2 \end{pmatrix} = (u_1)(av_1) + (u_2)(av_2) = a(u_1v_1 + u_2v_2)$$

and

$$a(\mathbf{u} \cdot \mathbf{v}) = a\left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right) = a(u_1v_1 + u_2v_2).$$

Therefore, we have $\mathbf{u} \cdot (a\mathbf{v}) = a(\mathbf{u} \cdot \mathbf{v})$ for all scalars a and vectors $\mathbf{u}$ and $\mathbf{v}$.

Turning to $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w})$ and $\mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$, we have

$$\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \cdot \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix} = u_1(v_1 + w_1) + u_2(v_2 + w_2)$$

and

$$\mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = u_1v_1 + u_2v_2 + u_1w_1 + u_2w_2 = u_1(v_1 + w_1) + u_2(v_2 + w_2).$$

As expected, we have $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$.

□

Important: For any scalar a and vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, we have

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u},$$
$$\mathbf{u} \cdot (a\mathbf{v}) = a(\mathbf{u} \cdot \mathbf{v}),$$
$$\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}.$$

As we noted on page 307, a function f with domain $\mathbb{R}^2$ is a **linear function** if $f(a\mathbf{v}) = af(\mathbf{v})$ for all scalars a and vectors $\mathbf{v}$, and $f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w})$ for all vectors $\mathbf{v}$ and $\mathbf{w}$. Together, the properties $\mathbf{u} \cdot (a\mathbf{v}) = a(\mathbf{u} \cdot \mathbf{v})$ and $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ tell us that for any vector $\mathbf{u}$, the function $f(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$ is a linear function. Combining this with the fact that $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$, we see that $f(\mathbf{v}) = \mathbf{v} \cdot \mathbf{u}$ is also a linear function. We care about this because so many relationships in a great many fields of study are linear relationships. After we get a little more experience with the dot product, we'll prove that *every linear function that maps vectors to scalars can be represented with a dot product.* So, by studying the properties of the dot product, we are studying the properties of linear functions that arise in a wide range of fields.

Sidenote: We say that the function $g(\mathbf{u}, \mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$ is **bilinear** because it is linear in both $\mathbf{u}$ and $\mathbf{v}$.

Problem 9.8: Show that for any vector $\mathbf{v}$, we have $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$.

Solution for Problem 9.8: Let $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, so that

$$\mathbf{v} \cdot \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = v_1^2 + v_2^2 = \|\mathbf{v}\|^2.$$

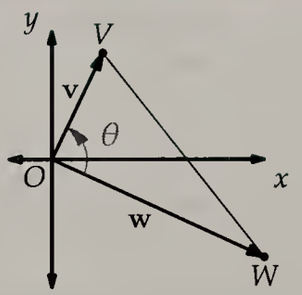
□

Important: For any vector $\mathbf{v}$, we have $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$.

We can use the basic principles we proved in Problems 9.7 and 9.8 to develop a geometric interpretation of the dot product.

Problem 9.9: Let θ be the angle between nonzero vectors $\mathbf{v}$ and $\mathbf{w}$. Show that $\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos\theta$.

Solution for Problem 9.9: The presence of $\|\mathbf{v}\|$, $\|\mathbf{w}\|$ and $\cos\theta$ suggest a geometric approach. At right, we have points V and W, where $\vec{V} = \mathbf{v}$ and $\vec{W} = \mathbf{w}$, and we let the origin be O. We therefore have $OV = \|\mathbf{v}\|$, $OW = \|\mathbf{w}\|$, and $WV = \left\|\vec{WV}\right\| = \|\mathbf{v} - \mathbf{w}\|$.



We try the Law of Cosines, since this will produce a term of the form $\|\mathbf{v}\| \|\mathbf{w}\| \cos\theta$. The Law of Cosines gives us $WV^2 = OV^2 + OW^2 - 2(OV)(OW)\cos\theta$, or

$$\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\|\mathbf{v}\| \|\mathbf{w}\| \cos\theta.$$

To introduce dot products, we apply the relationship we proved in Problem 9.8, which gives us $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$, $\|\mathbf{w}\|^2 = \mathbf{w} \cdot \mathbf{w}$, and $\|\mathbf{v} - \mathbf{w}\|^2 = (\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w})$, so we can write the Law of Cosines equation above as

$$(\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w}) = \mathbf{v} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w} - 2\|\mathbf{v}\| \|\mathbf{w}\| \cos\theta.$$

This equation doesn't have $\mathbf{v} \cdot \mathbf{w}$, yet. Applying the fact that the dot product is distributive, which we proved in Problem 9.7, we have

$$\begin{aligned}(\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w}) &= (\mathbf{v} - \mathbf{w}) \cdot \mathbf{v} - (\mathbf{v} - \mathbf{w}) \cdot \mathbf{w} \\ &= \mathbf{v} \cdot \mathbf{v} - \mathbf{w} \cdot \mathbf{v} - \mathbf{v} \cdot \mathbf{w} + \mathbf{w} \cdot \mathbf{w}.\end{aligned}$$

Since $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$, our Law of Cosines equation now is

$$\mathbf{v} \cdot \mathbf{v} - 2\mathbf{v} \cdot \mathbf{w} + \mathbf{w} \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w} - 2\|\mathbf{v}\| \|\mathbf{w}\| \cos\theta.$$

Simplifying this equation gives

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos\theta.$$

□

Important: If θ is the angle between nonzero vectors $\mathbf{v}$ and $\mathbf{w}$, then

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos\theta.$$

Rearranging this equation gives

$$\cos\theta = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|}.$$

We can use this relationship to find the angle between two vectors.

Problem 9.10: Find the angle between each pair of vectors below.

(a) $\begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix}$ and $\begin{pmatrix} 0 \\ -2 \end{pmatrix}$ (b) $\begin{pmatrix} 12 \\ -8 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ (c) $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -2 - \sqrt{3} \\ -1 + 2\sqrt{3} \end{pmatrix}$

Solution for Problem 9.10: In each of the following parts, let θ be the angle between the two given vectors.

(a) We have

$$\cos\theta = \frac{(-1)(0) + (\sqrt{3})(-2)}{\sqrt{(-1)^2 + \left(\sqrt{3}\right)^2} \cdot \sqrt{0^2 + 2^2}} = \frac{-2\sqrt{3}}{4} = -\frac{\sqrt{3}}{2},$$

so $\theta = 150°$.

(b) We have

$$\cos\theta = \frac{(12)(2) + (-8)(3)}{\sqrt{12^2 + (-8)^2} \cdot \sqrt{2^2 + 3^2}} = 0,$$

so $\theta = 90°$.

(c) We have

$$\cos\theta = \frac{(2)(-2 - \sqrt{3}) + (1)(-1 + 2\sqrt{3})}{\sqrt{2^2 + 1^2} \cdot \sqrt{(-2 - \sqrt{3})^2 + (-1 + 2\sqrt{3})^2}} = \frac{-5}{\sqrt{5} \cdot 2\sqrt{5}} = -\frac{1}{2},$$

so $\theta = 120°$.

□

Problem 9.11: We say that two vectors are **orthogonal** if their dot product is zero. Suppose we draw arrows representing two nonzero vectors from the same point. If these vectors are orthogonal, then how are the arrows that represent them related geometrically?

Solution for Problem 9.11: Let the vectors be $\mathbf{v}$ and $\mathbf{w}$. Since $\mathbf{v}$ and $\mathbf{w}$ are nonzero, we have $\|\mathbf{v}\|\,\|\mathbf{w}\|\cos\theta = 0$ if and only if $\cos\theta = 0$, which means the angle between the vectors is 90°. Therefore, the arrows representing the vectors are perpendicular if and only if the vectors are orthogonal.

So, we can interpret "orthogonal" geometrically to mean "perpendicular." Since these concepts are so closely related, they share a symbol. We can write that $\mathbf{v}$ and $\mathbf{w}$ are orthogonal as $\mathbf{v} \perp \mathbf{w}$. □

Important: If the dot product of two vectors is 0, then then the vectors are said to be **orthogonal** to each other. Geometrically speaking, two nonzero orthogonal vectors are perpendicular.

The vector $\mathbf{0}$ is orthogonal to all vectors.

Problem 9.12: If $\mathbf{p} \cdot \mathbf{q} = \mathbf{p} \cdot \mathbf{r}$ for a nonzero vector $\mathbf{p}$, then must we have $\mathbf{q} = \mathbf{r}$?

Solution for Problem 9.12: No. Rearranging the equation $\mathbf{p} \cdot \mathbf{q} = \mathbf{p} \cdot \mathbf{r}$ gives $\mathbf{p} \cdot \mathbf{q} - \mathbf{p} \cdot \mathbf{r} = 0$, which we can "factor" as $\mathbf{p} \cdot (\mathbf{q} - \mathbf{r}) = 0$, because the dot product is distributive. Therefore, if $\mathbf{p} \cdot \mathbf{q} = \mathbf{p} \cdot \mathbf{r}$, all we know is that the vector $\mathbf{q} - \mathbf{r}$ is orthogonal to $\mathbf{p}$. We cannot conclude that $\mathbf{q} = \mathbf{r}$.

Now that we know what to look for, we can quickly find an example for which $\mathbf{p} \cdot \mathbf{q} = \mathbf{p} \cdot \mathbf{r}$, but $\mathbf{q} \neq \mathbf{r}$. Let $\mathbf{p} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$. Then, we find x and y such that $\begin{pmatrix} 2 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = 0$. One such pair is $(x, y) = (1, 2)$. Therefore, if $\mathbf{q} - \mathbf{r} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, then we have $\mathbf{p} \cdot \mathbf{q} = \mathbf{p} \cdot \mathbf{r}$. We can choose any vector for $\mathbf{r}$ and solve for the corresponding $\mathbf{q}$. Specifically, suppose $\mathbf{r} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$, then we have $\mathbf{q} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \mathbf{r} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$, and indeed we have $\mathbf{p} \cdot \mathbf{q} = \mathbf{p} \cdot \mathbf{r} = 4$, but $\mathbf{q} \neq \mathbf{r}$. □

WARNING!! We cannot "divide" vectors from dot products. For example, if $\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{w}$, then we *cannot* deduce that $\mathbf{v}$ and $\mathbf{w}$ are the same.

Problem 9.13: Use the dot product to prove that

$$(x_1y_1 + x_2y_2)^2 \leq (x_1^2 + x_2^2)(y_1^2 + y_2^2)$$

for all real x_1, x_2, y_1, and y_2. When does equality hold?

Solution for Problem 9.13: The expression on the lesser side looks like the square of a dot product and the expression on the greater side looks like the product of the squares of vector norms. So, we let $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. Then, we have

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2$$

and

$$\|\mathbf{x}\|^2 \|\mathbf{y}\|^2 = (x_1^2 + x_2^2)(y_1^2 + y_2^2).$$

We also have $\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos\theta$, where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$. We must have $|\cos\theta| \leq 1$, so

$$\left| \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} \right| = |\cos\theta| \leq 1.$$

Squaring the left and right ends of this, multiplying both sides by $\|\mathbf{x}\|^2 \|\mathbf{y}\|^2$, and substituting our expressions for $\mathbf{x} \cdot \mathbf{y}$ and $\|\mathbf{x}\|^2 \|\mathbf{y}\|^2$ from above gives the desired

$$(x_1y_1 + x_2y_2)^2 \leq (x_1^2 + x_2^2)(y_1^2 + y_2^2).$$

(Note that we can perform the initial squaring because both sides of the inequality are nonnegative.)

We have equality if and only if $|\cos\theta| = 1$, which means $\theta = 0°$ or $\theta = 180°$. This occurs if and only if $\mathbf{x}$ and $\mathbf{y}$ are in the same direction, or in opposite directions. In either case, we must have $\mathbf{x} = c\mathbf{y}$ for some nonzero constant c. Conversely, if $\mathbf{x} = c\mathbf{y}$ for some nonzero constant c, then $\theta = 0°$ or $\theta = 180°$, in which case we have $|\cos\theta| = 1$, so equality holds. Finally, equality holds if either vector is the vector $\mathbf{0}$. Therefore, equality holds in the inequality if and only if $\mathbf{y}$ is $\mathbf{0}$ or if $\mathbf{x} = c\mathbf{y}$ for some constant c. (Letting $c = 0$ covers the case in which $\mathbf{x}$ is the zero vector.)

This result is called the **Cauchy-Schwarz Inequality**. □

We finish this section by returning to our claim back on page 308 that every linear function that maps vectors to scalars can be represented with a dot product.

Problem 9.14: We say that a function f with domain $\mathbb{R}^2$ is a **linear function** if

(i) $f(a\mathbf{v}) = af(\mathbf{v})$ for all scalars a and vectors $\mathbf{v}$, and

(ii) $f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w})$ for all vectors $\mathbf{v}$ and $\mathbf{w}$.

Suppose $g(\mathbf{w})$ is a linear function that maps vectors to scalars. That is, it takes a vector as an input and gives a scalar as an output. For example, $g(\mathbf{w}) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \cdot \mathbf{w}$ is one such function.

(a) Explain why $g(x\mathbf{i} + y\mathbf{j}) = xg(\mathbf{i}) + yg(\mathbf{j})$.

(b) Why does part (a) tell us that there exists some vector $\mathbf{v}$ such that $g(\mathbf{w}) = \mathbf{v} \cdot \mathbf{w}$ for all $\mathbf{w}$? Is this vector $\mathbf{v}$ unique?

Solution for Problem 9.14:

(a) Applying rule (ii) in our definition of a linear function, we have

$$g(x\mathbf{i} + y\mathbf{j}) = g(x\mathbf{i}) + g(y\mathbf{j}).$$

Applying rule (i) to each of the terms on the right gives

$$g(x\mathbf{i} + y\mathbf{j}) = g(x\mathbf{i}) + g(y\mathbf{j}) = xg(\mathbf{i}) + yg(\mathbf{j}).$$

(b) From part (a), we know that if $\mathbf{w} = w_1\mathbf{i} + w_2\mathbf{j}$, we have

$$g(\mathbf{w}) = g(w_1\mathbf{i} + w_2\mathbf{j}) = w_1 \cdot g(\mathbf{i}) + w_2 \cdot g(\mathbf{j}).$$

So, if we let $\mathbf{v} = \begin{pmatrix} g(\mathbf{i}) \\ g(\mathbf{j}) \end{pmatrix}$, then we have

$$g(\mathbf{w}) = w_1 \cdot g(\mathbf{i}) + w_2 \cdot g(\mathbf{j}) = \begin{pmatrix} g(\mathbf{i}) \\ g(\mathbf{j}) \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \mathbf{v} \cdot \mathbf{w}.$$

Since $g(\mathbf{i})$ and $g(\mathbf{j})$ are unique, this choice of $\mathbf{v}$ is unique. Therefore, we have shown that:

Important: Each linear function g that maps vectors to scalars can be defined as $g(\mathbf{w}) = \mathbf{v} \cdot \mathbf{w}$ for a unique vector $\mathbf{v}$.

This is why dot products are so important, and why you'll be seeing them throughout your future studies of science and mathematics.

□

Exercises

9.2.1 Evaluate each of the following

(a) $\begin{pmatrix} 5 \\ -3 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ 4 \end{pmatrix}$ (b) $\begin{pmatrix} -6 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -3 \end{pmatrix}$

9.2.2 Let $\mathbf{u} = 5\mathbf{i} - 12\mathbf{j}$ and $\mathbf{v} = 3\mathbf{i} + 4\mathbf{j}$. Find $(\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v})$.

9.2.3 Find all vectors that have 5 as one component and are orthogonal to $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

9.2.4 Which two pairs of the following 5 vectors are orthogonal: $\begin{pmatrix} 4 \\ -6 \end{pmatrix}, \begin{pmatrix} -3 \\ 4 \end{pmatrix}, \begin{pmatrix} 8 \\ 6 \end{pmatrix}, \begin{pmatrix} 6 \\ 9 \end{pmatrix}, \begin{pmatrix} 9 \\ 6 \end{pmatrix}$.

9.2.5 Find, to the nearest degree, the measures of the angles of $\triangle ABC$ if A is $(0,5)$, B is $(1,-3)$, and C is $(-5,4)$.

9.2.6 In this problem, we find another proof that the algebraic and geometric representations of the dot product are equivalent. Let $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$, and let ϕ be the angle between $\mathbf{v}$ and $\mathbf{w}$.

(a) Let $(r_\mathbf{v}, \theta_\mathbf{v})$ and $(r_\mathbf{w}, \theta_\mathbf{w})$ be the polar coordinates of the points whose rectangular coordinates are (v_1, v_2) and (w_1, w_2), respectively. Express $\|\mathbf{v}\| \|\mathbf{w}\| \cos\phi$ in terms of these polar coordinates.

(b) Show that your expression from part (a) equals $v_1w_1 + v_2w_2$.

9.3 Lines and Linear Dependence

In your study of algebra, you probably learned that if a, b, and c are constants, then the graph of the equation of $ax + by = c$ is a line. In this section, we revisit the algebra of lines in the Cartesian plane in the context of vectors.

Geometrically speaking, any two distinct points A and B determine a line $\overleftrightarrow{AB}$. We can think of this line as consisting of A together with all the points we can reach by going either in the direction of B or in the opposite direction (away from B). This leads us naturally to a description of a line using vectors.

As a reminder, parametric equations for x and y are equations that define the relationship between x and y through an intermediary variable. For example, consider the parametric equations $x = 3t$, $y = 2t$. We can write these parametric equations as a single parametric equation in terms of vectors:

$$\begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

When we graph such a parametric equation, we graph all points (x, y) such that there is a value of t for which (x, y) is a solution to the equation.

Problems

Problem 9.15: Let A be $(2, 3)$ and let B be $(-3, 1)$.

(a) Explain why $\overleftrightarrow{AB}$ is the graph of the parametric equation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} -5 \\ -2 \end{pmatrix}.$$

(b) Is this the only parametric equation of the form

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$$

whose graph is $\overleftrightarrow{AB}$?

Problem 9.16: Let $\mathbf{u}$ and $\mathbf{v}$ be vectors with $\mathbf{v} \neq \mathbf{0}$. Explain why the graph of the parametric equation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$$

fits our geometric description of a line.

Problem 9.17:

(a) Show that if the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$ contains (p, q), then the graph is the same as the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix} + s\mathbf{v}$.

(b) Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be nonzero vectors. Show that if the graphs of $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u}_1 + s\mathbf{v}_1$ and $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u}_2 + t\mathbf{v}_2$ both contain distinct points (x_a, y_a) and (x_b, y_b), then the graphs are the same and $\mathbf{v}_1 = k\mathbf{v}_2$ for some constant k.

Problem 9.18:

(a) Express the equation $2x - 3y = 6$ as a parametric equation of the form $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$.

(b) Show that the graph any equation of the form $ax + by = c$, where at least one of a and b is nonzero, is also the graph of a parametric equation of the form

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}.$$

(c) Show that the graph of any parametric equation of the form $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$, where $\mathbf{v}$ is nonzero, is the same as the graph of some equation of the form $ax + by = c$, where a, b, and c are constants.

(d) How is the vector $\mathbf{v}$ related to the slope of the line $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$?

Problem 9.19: The vector $a\mathbf{v} + b\mathbf{w}$, where a and b are scalars, is called a **linear combination** of $\mathbf{v}$ and $\mathbf{w}$. In this problem, we explore the graphs of all linear combinations of pairs of vectors. Specifically, we are interested in the graph of the parametric equation

$$\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w},$$

where both s and t are parameters.

(a) What is the graph if $\mathbf{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}$?

(b) What is the graph if $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$?

(c) What is the graph if $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$?

Problem 9.20: Let $\mathbf{v}$ and $\mathbf{w}$ be nonzero vectors. Show that the graph of the parametric equation $\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$, where both s and t are parameters, is a line if $\mathbf{v}$ is a scalar multiple of $\mathbf{w}$.

Problem 9.21: Let $\mathbf{v}$ and $\mathbf{w}$ be nonzero vectors such that $\mathbf{v}$ is not a scalar multiple of $\mathbf{w}$. In this problem, we show that every point (x, y) is in the graph of the parametric equation

$$\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w},$$

where s and t are parameters.

(a) What is the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{v} + t\mathbf{w}$? (Note that the coefficient of $\mathbf{v}$ is 1, not s.)

(b) What is the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = 2\mathbf{v} + t\mathbf{w}$?

(c) How are your graphs in parts (a) and (b) related?

(d) Show that every point (x, y) is in the graph of the parametric equation $\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$, where s and t are parameters.

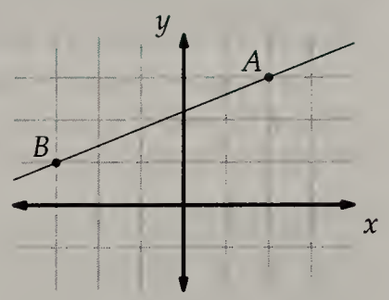
Problem 9.15: Let A be $(2,3)$ and let B be $(-3,1)$.

(a) Explain why $\overleftrightarrow{AB}$ is the graph of the parametric equation $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} -5 \\ -2 \end{pmatrix}$.

(b) Is this the only parametric equation of the form $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$ whose graph is this line?

Solution for Problem 9.15:

(a) In the introduction of this section, we described the line through A and B as being all the points we can reach by starting at point A and going either in the direction of B or in the opposite direction (away from B). The direction from A to B is the direction of $\vec{AB}$. We have $\vec{AB} = \begin{pmatrix} -3 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -5 \\ -2 \end{pmatrix}$. To get to any point on $\overleftrightarrow{AB}$, we start from $(2,3)$ and add some multiple of $\vec{AB}$. So, $\overleftrightarrow{AB}$ is the graph of



$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} -5 \\ -2 \end{pmatrix}.$$

(b) The equation we found in part (a) isn't the only parametric equation whose graph is $\overleftrightarrow{AB}$. We can choose any point on the line as our starting point, and any other point on the line to travel towards (or away from) to produce an equation for the line. For example, we could start from B and go towards or away from A. Then, we have have $\vec{BA} = -\vec{AB} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$, and the line is the graph of

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \end{pmatrix} + s \begin{pmatrix} 5 \\ 2 \end{pmatrix}.$$

Moreover, we can use any nonzero multiple of $\begin{pmatrix} 5 \\ 2 \end{pmatrix}$ as the direction vector. For example, if we let $s = 4k$ in the equation above, we have the parametric equation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \end{pmatrix} + k \begin{pmatrix} 20 \\ 8 \end{pmatrix},$$

which also has $\overleftrightarrow{AB}$ as its graph. So, the choice of $\mathbf{u}$ and $\mathbf{v}$ such that $\overleftrightarrow{AB}$ is the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$ is not unique. Vector $\mathbf{u}$ can correspond to any point on $\overleftrightarrow{AB}$, and $\mathbf{v}$ can be any nonzero multiple of $\begin{pmatrix} -5 \\ -2 \end{pmatrix}$.

□

Problem 9.16: Let $\mathbf{u}$ and $\mathbf{v}$ be vectors with $\mathbf{v} \neq \mathbf{0}$. Explain why the graph of the parametric equation $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$ fits our geometric description of a line given on page 313.

Solution for Problem 9.16: Let $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$. Then, the equation $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + t\mathbf{v}$ consists of (u_1, u_2) together with all points we can reach by starting from (u_1, u_2) and going in the direction of $\mathbf{v}$, or going in the direction opposite $\mathbf{v}$. This is precisely the description of a line given on page 313. □

The previous two problems are not a proof that the graph of the parametric equation $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$ is a line.

They explain that if we *define* a line in the Cartesian plane to be the set of points that satisfy this equation for some specific **u** and **v**, then this line fits our geometric description of a line.

Definition: If **u** and **v** are vectors with $\mathbf{v} \neq \mathbf{0}$, then the graph of the parametric equation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v},$$

where t is a parameter, is a **line**.

The vector **v** is called a **direction vector** of the line, and we'll sometimes describe a direction vector of a line as being "parallel" to the line. As suggested in our solution to Problem 9.15, the direction vector of a line is not unique. In the next problem, we investigate how direction vectors of the same line are related.

Problem 9.17:

(a) Show that if the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$ contains (p, q), then the graph is the same as the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix} + s\mathbf{v}$.

(b) Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be nonzero vectors. Show that if the graphs of $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u}_1 + s\mathbf{v}_1$ and $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u}_2 + t\mathbf{v}_2$ both contain distinct points (x_a, y_a) and (x_b, y_b), then the graphs are the same and $\mathbf{v}_1 = k\mathbf{v}_2$ for some constant k.

Solution for Problem 9.17:

(a) We defined a line algebraically using a point on the line and a direction. This part asks us to show that no matter which point on the line we choose, we'll get the same line.

Since (p, q) is in the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$, we have $\begin{pmatrix} p \\ q \end{pmatrix} = \mathbf{u} + t_1\mathbf{v}$ for some scalar t_1. Solving for **u** gives $\mathbf{u} = \begin{pmatrix} p \\ q \end{pmatrix} - t_1\mathbf{v}$. Substituting this into $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$ gives

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix} - t_1\mathbf{v} + t\mathbf{v} = \begin{pmatrix} p \\ q \end{pmatrix} + (t - t_1)\mathbf{v}. \tag{9.1}$$

To see that Equation (9.1) has the same graph as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix} + s\mathbf{v}, \tag{9.2}$$

we note that for any choice of s in (9.2), letting $t = s + t_1$ in (9.1) gives the same $\begin{pmatrix} x \\ y \end{pmatrix}$. Conversely, for any choice of t in (9.1), choosing $s = t - t_1$ in (9.2) gives the same $\begin{pmatrix} x \\ y \end{pmatrix}$. Therefore, every point in the graph of (9.1) is in the graph of (9.2), and vice versa, so the graphs are the same.

(b) Geometrically speaking, this problem is a fancy way way of saying "two points determine a line." That is, if two lines have two points in common, then the lines must be the same. Let's show that our algebraic definition of a line satisfies this principle.

We start by using part (a) to note that the graphs of the two given parametric equations are also the

graphs of

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_a \\ y_a \end{pmatrix} + s\mathbf{v}_1, \tag{9.3}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_a \\ y_a \end{pmatrix} + t\mathbf{v}_2, \tag{9.4}$$

respectively. Since (x_b, y_b) is in both graphs and is distinct from (x_a, y_a), there are nonzero scalars s_b and t_b such that

$$\begin{pmatrix} x_b \\ y_b \end{pmatrix} = \begin{pmatrix} x_a \\ y_a \end{pmatrix} + s_b\mathbf{v}_1,$$

$$\begin{pmatrix} x_b \\ y_b \end{pmatrix} = \begin{pmatrix} x_a \\ y_a \end{pmatrix} + t_b\mathbf{v}_2,$$

Taking the difference of these equations gives $s_b\mathbf{v}_1 - t_b\mathbf{v}_2 = \mathbf{0}$, so $\mathbf{v}_1 = \frac{t_b}{s_b}\mathbf{v}_2$. Letting $k = \frac{t_b}{s_b}$ gives us $\mathbf{v}_1 = k\mathbf{v}_2$, as desired.

We now must show that graphs of (9.3) and (9.4) are the same. Substituting $\mathbf{v}_1 = k\mathbf{v}_2$ into (9.3) gives us

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_a \\ y_a \end{pmatrix} + sk\mathbf{v}_2, \tag{9.5}$$

where s is a parameter and k is constant. All we have left is showing that this equation has the same graph as (9.4). Fortunately, that's easy. For any choice of t in (9.4), letting $s = \frac{t}{k}$ in (9.5) produces the same $\begin{pmatrix} x \\ y \end{pmatrix}$. Conversely, for any choice of s in (9.5), choosing $t = sk$ in (9.4) produces the same $\begin{pmatrix} x \\ y \end{pmatrix}$. Therefore, Equations (9.4) and (9.5) have the same graph.

□

Problem 9.18:

(a) Express the equation $2x - 3y = 6$ as a parametric equation of the form $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$.

(b) Show that the graph any equation of the form $ax + by = c$, where at least one of a and b is nonzero, is also the graph of a parametric equation of the form $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$.

(c) Show that the graph of any parametric equation of the form $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$, where $\mathbf{v}$ is nonzero, is the same as the graph of some equation of the form $ax + by = c$, where a, b, and c are constants.

(d) How is the vector $\mathbf{v}$ related to the slope of the line $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$?

Solution for Problem 9.18:

(a) Solving for x gives $x = \frac{3}{2}y + 3$. Letting $y = t$ gives us the parametric equations $x = \frac{3}{2}t + 3$, $y = t$, or $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix} + t\begin{pmatrix} 3/2 \\ 1 \end{pmatrix}$.

(b) If $a \neq 0$, we can use part (a) as a guide. Solving $ax + by = c$ for x gives $x = -\frac{b}{a}y + \frac{c}{a}$. We can take $t = y$ as the parameter, and we have

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c/a \\ 0 \end{pmatrix} + t\begin{pmatrix} -b/a \\ 1 \end{pmatrix}.$$

If $a = 0$, we can simply solve for y, since a and b cannot both be 0. We have $y = -\frac{a}{b}x + \frac{c}{b} = \frac{c}{b}$, and taking $t = x$ as the parameter gives

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ c/b \end{pmatrix} + t \begin{pmatrix} 1 \\ -a/b \end{pmatrix}.$$

(c) Let $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. Then, we have

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + t \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} u_1 + tv_1 \\ u_2 + tv_2 \end{pmatrix},$$

so $x = u_1 + tv_1$, $y = u_2 + tv_2$. To produce an equation of the form $ax + by = c$, we eliminate the parameter t. Multiplying $x = u_1 + tv_1$ by v_2 and multiplying $y = u_2 + tv_2$ by v_1 gives

$$v_2 x = u_1 v_2 + t v_1 v_2,$$
$$v_1 y = u_2 v_1 + t v_1 v_2.$$

Subtracting the second equation from the first eliminates the parameter t and leaves us with

$$v_2 x - v_1 y = u_1 v_2 - u_2 v_1.$$

Letting $a = v_2$, $b = -v_1$, and $c = u_1 v_2 - u_2 v_1$, we see that this equation is in the form $ax + by = c$.

Unfortunately, we're not quite finished. Our work above only tells us that every point in the graph of

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v} \tag{9.6}$$

is in the graph of

$$v_2 x - v_1 y = u_1 v_2 - u_2 v_1. \tag{9.7}$$

It does *not* tell us that every point in the graph of (9.7) is in the graph of (9.6). Fortunately, we've already done most of the work to go the other direction. Part (a) tells us that the graph of $v_2 x - v_1 y = u_1 v_2 - u_2 v_1$ is the graph of a parametric equation of the form

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u}' + t'\mathbf{v}' \tag{9.8}$$

for some vectors $\mathbf{u}'$ and $\mathbf{v}'$. We showed in part (b) of Problem 9.17 that if the graphs of two equations of this form share two points, then the graphs are the same. We know that every point in the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$ (and this graph definitely has more than one point because $\mathbf{v}$ is nonzero) is in the graph of (9.7), and therefore is in the graph of (9.8). Therefore, we conclude that the graphs of (9.6) and (9.8) are the same, which means that the graphs of (9.6) and (9.7) are the same.

(d) In part (c), we saw that the parametric equation $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$ is has the same graph as $v_2 x - v_1 y = u_1 v_2 - u_2 v_1$. We then have two cases to consider:

Case 1: $v_1 = 0$. If $v_1 = 0$, then $v_2 x - v_1 y = u_1 v_2 - u_2 v_1$ is simply $v_2 x = u_1 v_2$. Dividing by v_2 ($v_2 \neq 0$ because $\mathbf{v} \neq \mathbf{0}$), we have $x = u_1$. The graph of this equation is a vertical line with undefined slope. Looking back at the equation $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$, we see that when $v_1 = 0$, the direction vector is $\begin{pmatrix} 0 \\ v_2 \end{pmatrix}$, which is indeed a "vertical" vector. Therefore, a line with undefined slope has a "vertical" direction vector, as expected.

Case 2: $v_1 \neq 0$. If $v_1 \neq 0$, then we can solve $v_2 x - v_1 y = u_1 v_2 - u_2 v_1$ for y to get $y = \frac{v_2}{v_1}x - \frac{u_1 v_2}{v_1} + u_2$. The slope of this line is $\frac{v_2}{v_1}$. So, the slope of a line is the ratio of the vertical component of the direction vector to the horizontal component of the direction vector—this ratio is the traditional "rise over run" description of slope! □

We have therefore seen that our definition of a line with vectors fits nicely with the "graph of $ax + by = c$" understanding of lines in the Cartesian plane from basic algebra, where the direction vector of a line plays the role of "slope."

Problem 9.19: The vector $a\mathbf{v} + b\mathbf{w}$, where a and b are scalars, is called a **linear combination** of $\mathbf{v}$ and $\mathbf{w}$. In this problem, we explore the graphs of all linear combinations of pairs of vectors. Specifically, we are interested in the graph of the parametric equation

$$\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w},$$

where both s and t are parameters.

(a) What is the graph if $\mathbf{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}$?

(b) What is the graph if $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$?

(c) What is the graph if $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$?

Solution for Problem 9.19:

(a) We have

$$\begin{pmatrix} x \\ y \end{pmatrix} = s\begin{pmatrix} 2 \\ 1 \end{pmatrix} + t\begin{pmatrix} 6 \\ 3 \end{pmatrix} = \begin{pmatrix} 2s + 6t \\ s + 3t \end{pmatrix}.$$

The first component in $\begin{pmatrix} 2s + 6t \\ s + 3t \end{pmatrix}$ is double the second one, so $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2s + 6t \\ s + 3t \end{pmatrix}$ gives us $x = 2y$. Therefore, we see that the graph is a line through the origin.

(b) We have

$$\begin{pmatrix} x \\ y \end{pmatrix} = s\begin{pmatrix} 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} s \\ t \end{pmatrix}.$$

We can choose s and t to be anything, so the graph is the entire Cartesian plane.

(c) We have

$$\begin{pmatrix} x \\ y \end{pmatrix} = s\begin{pmatrix} 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} s - 2t \\ t \end{pmatrix}.$$

To show that every point (a, b) is in the graph, we let $s = a + 2b$ and $t = b$ in the parametric equation above, which gives $(x, y) = (a, b)$. Therefore, once again all possible (x, y) are in the graph, so the graph is the entire Cartesian plane.

□

The graph in part (a) is a line, while the graphs in parts (b) and (c) are the entire Cartesian plane. One key difference we notice in the two parts is that $\mathbf{v}$ is a scalar multiple of $\mathbf{w}$ in part (a), but not in parts (b) or (c). Let's investigate further.

Problem 9.20: Let $\mathbf{v}$ and $\mathbf{w}$ be nonzero vectors. Show that the graph of the parametric equation

$$\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w},$$

where both s and t are parameters, is a line if $\mathbf{v}$ is a scalar multiple of $\mathbf{w}$.

Solution for Problem 9.20: If $\mathbf{v} = k\mathbf{w}$ for some scalar k, then we have

$$\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w} = ks\mathbf{w} + t\mathbf{w} = (ks + t)\mathbf{w}.$$

The graph of this equation is indeed a line with direction vector $\mathbf{w}$. □

Now, let's take a closer look at what happens when $\mathbf{v}$ is not a multiple of $\mathbf{w}$.

Problem 9.21: Let $\mathbf{v}$ and $\mathbf{w}$ be nonzero vectors such that $\mathbf{v}$ is not a scalar multiple of $\mathbf{w}$. In this problem, we show that every point (x, y) is in the graph of the parametric equation $\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$, where s and t are parameters.

(a) What is the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{v} + t\mathbf{w}$? (Note that the coefficient of $\mathbf{v}$ is 1, not s.)

(b) What is the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = 2\mathbf{v} + t\mathbf{w}$?

(c) How are your graphs in parts (a) and (b) related?

(d) Show that every point (x, y) is in the graph of the parametric equation $\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$, where s and t are parameters.

Solution for Problem 9.21: Let $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$.

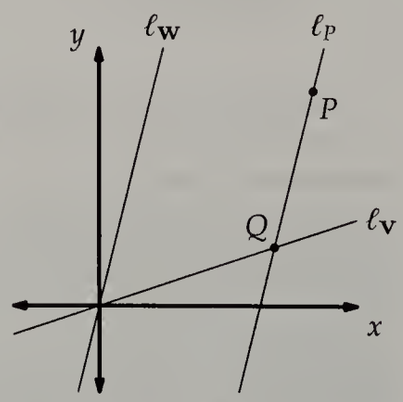
(a) By our definition of a line, the graph is the line through (v_1, v_2) in the direction of $\mathbf{w}$.

(b) By our definition of a line, the graph is the line through $(2v_1, 2v_2)$ in the direction of $\mathbf{w}$.

(c) The direction vectors of the lines in parts (a) and (b) are the same, so we suspect the two graphs do not intersect. The lines intersect at (x, y) if there are constants t_1 and t_2 such that $\mathbf{v} + t_1\mathbf{w} = 2\mathbf{v} + t_2\mathbf{w}$. Solving for $\mathbf{v}$ gives $\mathbf{v} = (t_1 - t_2)\mathbf{w}$. But we are told that $\mathbf{v}$ is not a scalar multiple of $\mathbf{w}$, so this is impossible. Therefore, the lines of parts (a) and (b) do not intersect.

(d) Continuing with our arguments from parts (a) and (b), for each possible value of s, the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$ is a line through (sv_1, sv_2) with direction vector $\mathbf{w}$. Continuing our reasoning from part (c), we see that all of these lines are parallel. So, the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$ is a set of parallel lines. We suspect that any point (x_0, y_0) is on one of these lines, but we have to prove it.

We'll start by developing some geometric intuition for why any (x_0, y_0) must be on one of these parallel lines. At right, point P is (x_0, y_0); line $\ell_{\mathbf{v}}$ is the line through the origin with direction $\mathbf{v}$, and line $\ell_{\mathbf{w}}$ is the line through the origin with direction $\mathbf{w}$. To show that P is on one of the set of parallel lines that is the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$, we draw line ℓ_P through P parallel to $\ell_{\mathbf{w}}$. Since $\mathbf{v}$ is not a constant multiple of $\mathbf{w}$, lines $\ell_{\mathbf{v}}$ and ℓ_P are not parallel and must therefore intersect at a point we'll call Q. Since Q is on $\ell_{\mathbf{v}}$, we have $\vec{Q} = s\mathbf{v}$ for some scalar s. Since $\ell_P \parallel \ell_{\mathbf{w}}$, we have $\vec{QP} = t\mathbf{w}$ for some scalar t. Therefore, $\begin{pmatrix} x \\ y \end{pmatrix} = \vec{P} = s\mathbf{v} + t\mathbf{w}$.



While that gives us some geometric intuition for why (x_0, y_0) must be on one of these parallel lines, we'd like our proof to rest on the same definition of a line that we've been using throughout this section. In

Problem 9.18, we showed that the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{v} + t\mathbf{w}$ is also the graph of the equation

$$w_2x - w_1y = w_2v_1 - w_1v_2.$$

Note that $(x, y) = (v_1, v_2)$ satisfies this equation, which is expected, since (v_1, v_2) is on the line. This equation describes the line in part (a). In part (b), we considered the line $\begin{pmatrix} x \\ y \end{pmatrix} = 2\mathbf{v} + t\mathbf{w}$. Again applying the result from Problem 9.18, this parametric equation can also be written as

$$w_2x - w_1y = w_2(2v_1) - w_1(2v_2) = 2(w_2v_1 - w_1v_2).$$

Similarly, for any specific value of s, the parametric equation $\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$, where t is a parameter and s is fixed, can also be expressed as

$$w_2x - w_1y = w_2(sv_1) - w_1(sv_2) = s(w_2v_1 - w_1v_2). \tag{9.9}$$

We wish to show that for any point (x_0, y_0), we can choose a value of s such that (x_0, y_0) is on graph of Equation (9.9). We know that (x_0, y_0) lies on

$$w_2x - w_1y = w_2x_0 - w_1y_0,$$

since this equation is satisfied by $(x, y) = (x_0, y_0)$. So, we choose $s = \frac{w_2x_0 - w_1y_0}{w_2v_1 - w_1v_2}$ in Equation (9.9):

$$w_2x - w_1y = s(w_2v_1 - w_1v_2) = \frac{w_2x_0 - w_1y_0}{w_2v_1 - w_1v_2}(w_2v_1 - w_1v_2) = w_2x_0 - w_1y_0.$$

Therefore, the point (x_0, y_0) is indeed on one of the set of parallel lines described by $\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$. Since this argument holds for any point (x_0, y_0), the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$ is the entire Cartesian plane.

Our argument appears to show that the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$ is the entire Cartesian plane, but in Problem 9.19, we saw that the graph is just a line if $\mathbf{v}$ is a constant multiple of $\mathbf{w}$. So, where does our argument fail if $\mathbf{v} = k\mathbf{w}$?

The only step in our argument that might fail is when we let $s = \frac{w_2x_0 - w_1y_0}{w_2v_1 - w_1v_2}$. What if the denominator, $w_2v_1 - w_1v_2$, is 0? Then, there's no such s. If $\mathbf{v} = k\mathbf{w}$, then $v_1 = kw_1$ and $v_2 = kw_2$, so we find that $w_2v_1 - w_1v_2 = kw_1w_2 - kw_1w_2 = 0$. This is why the argument fails when $\mathbf{v} = k\mathbf{w}$. But is this the only case in which $w_2v_1 - w_1v_2 = 0$?

Fortunately, yes, it is. We write $w_2v_1 - w_1v_2 = 0$ as $w_2v_1 = w_1v_2$. If w_1 and w_2 are nonzero, we can write this equation as $\frac{v_1}{w_1} = \frac{v_2}{w_2}$. Setting these equal to k gives us $v_1 = kw_1$ and $v_2 = kw_2$, so $\mathbf{v} = k\mathbf{w}$. If w_1 is 0, then $w_2 \neq 0$, since $\mathbf{w} \neq \mathbf{0}$. Then, $w_2v_1 = w_1v_2$ tells us that $v_1 = 0$ and we again have $\mathbf{v} = k\mathbf{w}$ for some k, namely $k = \frac{v_2}{w_2}$. Similarly, if $w_2 = 0$, we have $\mathbf{v} = \frac{v_1}{w_1}\mathbf{w}$.

Putting this all together, we see that $w_2v_1 - w_1v_2 = 0$ if and only if $\mathbf{v}$ is a multiple of $\mathbf{w}$. Therefore, we have $w_2v_1 - w_1v_2 \neq 0$ if and only if $\mathbf{v}$ is not a multiple of $\mathbf{w}$, and our argument above shows that the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$ is the entire Cartesian plane if $w_2v_1 - w_1v_2 \neq 0$. So, the graph is the entire Cartesian plane whenever $\mathbf{v}$ is not a multiple of $\mathbf{w}$.

□

In the last two problems, we investigated the graph of the parametric equation

$$\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{v} + t\mathbf{w},$$

where s and t are parameters. We found that the shape of the graph depends on whether or not $\mathbf{v}$ is a multiple of $\mathbf{w}$. This distinction is so important that we have a special name for it. We say that $\mathbf{v}$ and $\mathbf{w}$ are **linearly dependent** if one is a multiple of the other, and that $\mathbf{v}$ and $\mathbf{w}$ are **linearly independent** if neither is a multiple of the other.

As we continue our study of vectors, we'll want to expand our definition of "linearly dependent" and "linearly independent" to describe groups of more than 2 vectors. Here's how we do so:

Definition: We say that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are **linearly dependent** if there are scalars $a_1, a_2, \ldots, a_n$, not all 0, such that

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n = \mathbf{0}.$$

If the only scalars that satisfy this equation are $a_1 = a_2 = \cdots = a_n = 0$, then the vectors are **linearly independent**.

To see why this definition is equivalent to our initial definition of "linearly dependent" when we have only two vectors, suppose that $\mathbf{v}$ and $\mathbf{w}$ are nonzero and linearly dependent by our first definition. Then, we must have $\mathbf{v} = k\mathbf{w}$. Letting $a_1 = 1$, $a_2 = -k$ gives us $a_1\mathbf{v} + a_2\mathbf{w} = \mathbf{0}$, which shows that $\mathbf{v}$ and $\mathbf{w}$ are linearly dependent by our new definition. If we start instead with the fact that $\mathbf{v}$ and $\mathbf{w}$ are nonzero and linearly dependent by our new definition, then there must be scalars a_1 and a_2, not both zero, such that $a_1\mathbf{v} + a_2\mathbf{w} = \mathbf{0}$. We can't have $a_1 = 0$, since this would mean $a_2\mathbf{w} = \mathbf{0}$, but $\mathbf{w}$ is nonzero and we cannot have $a_1 = a_2 = 0$. Since $a_1 \neq 0$, we can solve $a_1\mathbf{v} + a_2\mathbf{w} = \mathbf{0}$ for $\mathbf{v}$ to find $\mathbf{v} = -\frac{a_2}{a_1}\mathbf{w}$, which satisfies our first definition of linearly dependent.

We also have to make sure the definitions are equivalent if either vector is $\mathbf{0}$. If $\mathbf{v} = \mathbf{0}$, then $\mathbf{w} = 0\mathbf{v}$, so $\mathbf{v}$ and $\mathbf{w}$ are linearly dependent by our first definition. Also, we have $1\mathbf{v} + 0\mathbf{w} = \mathbf{0}$ in this case, so the vectors are linearly dependent by our second definition, as well. Similarly, the definitions are equivalent if $\mathbf{w} = \mathbf{0}$.

Therefore, the two definitions given for two vectors being linearly dependent are equivalent.

Important: Two vectors are linearly dependent if and only if one is a multiple of the other. Otherwise, they are linearly independent.

Exercises

9.3.1 Find an equation of the form $ax + by = c$ that has the same graph as the parametric equation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix} + s\begin{pmatrix} -4 \\ 5 \end{pmatrix}.$$

9.3.2 Find a parametric equation of the form $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$ that has the same graph as $3x - y = 4$.

9.3.3 Find all pairs of linearly dependent vectors in the list $\begin{pmatrix} 4 \\ -6 \end{pmatrix}, \begin{pmatrix} -9 \\ 6 \end{pmatrix}, \begin{pmatrix} 12 \\ 8 \end{pmatrix}, \begin{pmatrix} -6 \\ 9 \end{pmatrix}, \begin{pmatrix} 9 \\ 6 \end{pmatrix}$.

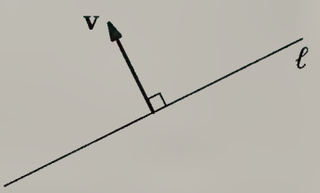
9.3.4

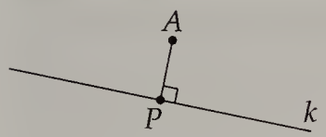
(a) Show that if $\mathbf{u}$ and $\mathbf{v}$ are linearly independent, then for each vector $\mathbf{w}$, then there is a unique pair of scalars a and b such that $\mathbf{w} = a\mathbf{u} + b\mathbf{v}$.

(b) Show that if there is more than one pair of scalars a and b such that $a\mathbf{u} + b\mathbf{v} = \mathbf{w}$ for three given vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, then $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent.

9.4 Projections

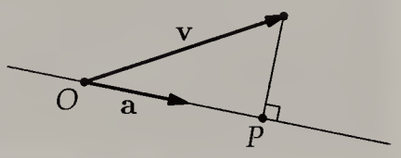
A vector is **normal** to a line if it is orthogonal to any vector in the direction of that line. So, for example, $\mathbf{j}$ is normal to the graph of the line $y = 3$, the vector $\mathbf{i} + \mathbf{j}$ is normal to the graph of $x + y = 7$, and $\mathbf{v}$ is normal to line ℓ in the diagram at right.



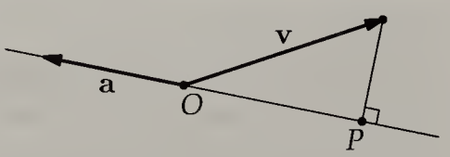


The **projection** of a point onto a line is the foot of the perpendicular from the point to the line. So, in the diagram at left, P is the projection of point A onto k.

We can extend the concept of projection to vectors. To find the projection of $\mathbf{v}$ onto a nonzero vector $\mathbf{a}$, we draw $\mathbf{v}$ from a point O, and then draw a line through O with direction vector $\mathbf{a}$. We then let P be the projection of the head of $\mathbf{v}$ onto this line. We call $\vec{OP}$ the projection of $\mathbf{v}$ onto $\mathbf{a}$. (Note that O does not have to be the origin in this definition!) We denote the projection of $\mathbf{v}$ onto $\mathbf{a}$ as $\text{proj}_{\mathbf{a}}\mathbf{v}$, so $\text{proj}_{\mathbf{a}}\mathbf{v} = \vec{OP}$.



One reason we draw a line in the direction of $\mathbf{a}$ rather than just drawing $\mathbf{a}$ or a ray in the direction of $\mathbf{a}$ is shown at right. Here, the projection of the head of $\mathbf{v}$ onto the line is on the opposite side of O from $\mathbf{a}$. Even though $\vec{OP}$ is in the opposite direction from $\mathbf{a}$, we call $\vec{OP}$ the projection of $\mathbf{v}$ onto $\mathbf{a}$.



Problems

Problem 9.22:

(a) Find a vector that is normal to the graph of $3x + y = 0$.

(b) Find a vector that is normal to the graph of $3x + y + 10 = 0$.

(c) Let a, b, and c be constants such that at least one of a and b is nonzero. Find a vector in terms of a, b, and c that is normal the graph of $ax + by + c = 0$.

Problem 9.23:

(a) Show that if $\mathbf{n}$ is normal to line ℓ, then $k\mathbf{n}$ is normal to line ℓ for any constant k.

(b) Show that if $\mathbf{n}_1$ and $\mathbf{n}_2$ are nonzero vectors that are normal to the same line ℓ, then $\mathbf{n}_1$ and $\mathbf{n}_2$ are linearly dependent.

Problem 9.24: Let $\mathbf{v} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

(a) Find $\text{proj}_{\mathbf{i}}\mathbf{v}$ and $\text{proj}_{\mathbf{j}}\mathbf{v}$.

(b) Find the projection of $\mathbf{v}$ onto $2\mathbf{i}$.

(c) Find the projection of $\mathbf{v}$ onto $-776\mathbf{j}$.

Problem 9.25: In this problem, we find the projection of $\mathbf{v} = \begin{pmatrix} 4\sqrt{3} \\ 4 \end{pmatrix}$ onto the vector $\mathbf{w} = \begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}$. Let θ be the angle between $\mathbf{v}$ and $\mathbf{w}$, let O be the origin, and let V and W be points such that $\vec{V} = \mathbf{v}$ and $\vec{W} = \mathbf{w}$.

(a) Let P be the foot of the perpendicular from V to $\overleftrightarrow{OW}$. What is OP in terms of $\|\mathbf{v}\|$, $\|\mathbf{w}\|$, and θ?

(b) Find $\text{proj}_{\mathbf{w}}\mathbf{v}$.

Problem 9.26: Find $\text{proj}_{\mathbf{w}}\mathbf{v}$ for any two vectors $\mathbf{v}$ and $\mathbf{w}$ in terms of $\mathbf{v} \cdot \mathbf{w}$, $\|\mathbf{w}\|$, and $\mathbf{w}$.

Problem 9.27: Let $\mathbf{a} = 4\mathbf{i} + 3\mathbf{j}$ and $\mathbf{b} = -\mathbf{i} + 3\mathbf{j}$, and let $\mathbf{p} = \text{proj}_{\mathbf{b}}\mathbf{a}$.

(a) Find $\mathbf{p}$.

(b) Let $\mathbf{n} = \mathbf{a} - \mathbf{p}$. Find $\mathbf{b} \cdot \mathbf{n}$.

(c) Explain the geometric significance of your result in part (b).

Problem 9.28: The distance between a point and a line is the length of the shortest possible segment from the point to the line.

(a) Why does the distance between a point and a line equal the distance between the point and the projection of the point onto the line?

(b) Find the distance between $(4, -2)$ and the graph of $3x - 2y - 7 = 0$. **Hints:** 45, 134

(c) Generalize your results in part (b) by finding an expression for the distance between the point (x_0, y_0) and the line $ax + by + c = 0$, where a, b, c, x_0, and y_0 are constants.

Problem 9.22:

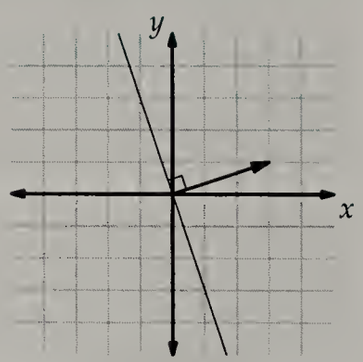
(a) Find a vector that is normal to the graph of $3x + y = 0$.

(b) Find a vector that is normal to the graph of $3x + y + 10 = 0$.

(c) Let a, b, and c be constants such that at least one of a and b is nonzero. Find a vector in terms of a, b, and c that is normal the graph of $ax + by + c = 0$.

Solution for Problem 9.22:

(a) First, we'll get an intuitive understanding what's going on here by using our ideas of "slope" and "perpendicular" from basic algebra. Let the graph of $3x + y = 0$ be line ℓ. Writing the equation as $y = -3x$, we see that the line has slope -3. Therefore, any line perpendicular to ℓ has slope $1/3$. So, a line through the origin perpendicular to ℓ must pass through $(3, 1)$. Therefore, one vector that should be normal to ℓ is $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$.



This isn't an entirely satisfying solution; it completely avoids the question of why the slope of the line perpendicular to ℓ is $1/3$. Instead of relying on "slope" and "perpendicular," let's use the definition of a normal vector. A normal vector must be orthogonal to any direction vector of the line. So, let's try to find a direction vector of the line. If (x_1, y_1) and (x_2, y_2) are two different points on the line, then a vector from one to the other, such as $\begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}$, is in the direction of the line. Since (x_1, y_1) and (x_2, y_2) are on the line, we have

$$3x_1 + y_1 = 0,$$
$$3x_2 + y_2 = 0.$$

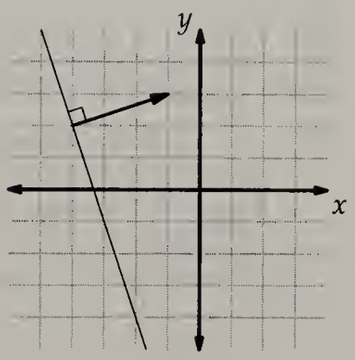
Subtracting the first equation from the second, we have $3(x_2 - x_1) + 1(y_2 - y_1) = 0$. Writing the left side as a dot product, out pops the normal vector:

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix} = 0.$$

This tells us that $\begin{pmatrix}3\\1\end{pmatrix}$ is orthogonal to any vector between two points in the line, so it is normal to the line.

The vector $\begin{pmatrix}3\\1\end{pmatrix}$ isn't the only vector normal to ℓ. As we'll investigate in Problem 9.23, any scalar multiple of $\begin{pmatrix}3\\1\end{pmatrix}$ is also normal to ℓ, and any vector that is normal to ℓ must be a scalar multiple of $\begin{pmatrix}3\\1\end{pmatrix}$ (including negative multiples and **0**).

(b) In part (a), the coefficients of x and y gave us components of a vector normal to the line. Maybe that just happens because the line $3x + y = 0$ goes through the origin. If we let (x_1, y_1) and (x_2, y_2) be points on the line $3x + y + 10 = 0$, we have



$$3x_1 + y_1 + 10 = 0,$$
$$3x_2 + y_2 + 10 = 0.$$

Subtracting the first from the last gives us $3(x_2 - x_1) + 1(y_2 - y_1) = 0$, and again we have

$$\begin{pmatrix}3\\1\end{pmatrix} \cdot \begin{pmatrix}x_2 - x_1\\y_2 - y_1\end{pmatrix} = 0.$$

So, the vector $\begin{pmatrix}3\\1\end{pmatrix}$ is orthogonal to any vector between two points on the line, and is therefore normal to the line.

(c) The first two parts give us a guide. We let (x_1, y_1) and (x_2, y_2) be on the line, so

$$ax_1 + by_1 + c = 0,$$
$$ax_2 + by_2 + c = 0.$$

Subtracting the first equation from the second gives $a(x_2 - x_1) + b(y_2 - y_1) = 0$, so

$$\begin{pmatrix}a\\b\end{pmatrix} \cdot \begin{pmatrix}x_2 - x_1\\y_2 - y_1\end{pmatrix} = 0,$$

which means that $\begin{pmatrix}a\\b\end{pmatrix}$ is orthogonal to any vector between two points on the line. Therefore, $\begin{pmatrix}a\\b\end{pmatrix}$ is normal to the line.

Important: The vector $\begin{pmatrix}a\\b\end{pmatrix}$ is normal to the graph of $ax + by + c = 0$.

□

Sidenote: While we'll be working with normal vectors only in the context of lines and planes in this book, you'll be seeing them again in the context of curves and curved surfaces when you study calculus, physics, and other subjects. We say that a vector is **normal** to a curve at a certain point on the curve if the vector is normal to the line through that point that is tangent to the curve. In three dimensions, a vector that is normal to a surface at a point is normal to the plane through that point that is tangent to the surface.

Problem 9.23:

(a) Show that if $\mathbf{n}$ is normal to line ℓ, then $k\mathbf{n}$ is normal to line ℓ for any constant k.

(b) Show that if $\mathbf{n}_1$ and $\mathbf{n}_2$ are nonzero vectors that are normal to the same line ℓ, then $\mathbf{n}_1$ and $\mathbf{n}_2$ are linearly dependent.

Solution for Problem 9.23:

(a) Let $\mathbf{d}$ be a direction vector of ℓ, so $\mathbf{n} \cdot \mathbf{d} = 0$. We showed in Problem 9.17 that any vector in the direction of line ℓ is of the form $c\mathbf{d}$ for some scalar c. For any scalar k, we have $(k\mathbf{n}) \cdot (c\mathbf{d}) = kc(\mathbf{n} \cdot \mathbf{d}) = 0$, so the vector $k\mathbf{n}$ is orthogonal to every vector in the direction of ℓ, which means $k\mathbf{n}$ is normal to ℓ.

(b) We'll first think about the geometry of the problem, and then provide an algebraic proof. We can think of a vector that is normal to ℓ as being the direction vector of a line perpendicular to ℓ. For any point, there is only one line through the point perpendicular to ℓ, and we showed in Problem 9.17 that any two possible direction vectors of the same line must be linearly dependent. Moreover, all lines perpendicular to ℓ are parallel to each other, so any two possible direction vectors of lines normal to ℓ are linearly dependent.

To prove algebraically that $\mathbf{n}_1$ and $\mathbf{n}_2$ are linearly dependent, we will show that it is impossible for them to be linearly independent. We start by noting that if $\mathbf{d}$ is a direction vector of ℓ, then we have $\mathbf{n}_1 \cdot \mathbf{d} = 0$ and $\mathbf{n}_2 \cdot \mathbf{d} = 0$. If $\mathbf{n}_1$ and $\mathbf{n}_2$ are linearly independent, the graph of $\begin{pmatrix} x \\ y \end{pmatrix} = s\mathbf{n}_1 + t\mathbf{n}_2$ passes through every point (x, y). Specifically, this means that there are some constants a_1 and a_2 such that $\mathbf{d} = a_1\mathbf{n}_1 + a_2\mathbf{n}_2$. Multiplying $\mathbf{n}_1 \cdot \mathbf{d} = 0$ and $\mathbf{n}_2 \cdot \mathbf{d} = 0$ by a_1 and a_2, respectively, gives

$$a_1\mathbf{n}_1 \cdot \mathbf{d} = 0,$$
$$a_2\mathbf{n}_2 \cdot \mathbf{d} = 0.$$

Adding these gives $a_1\mathbf{n}_1 \cdot \mathbf{d} + a_2\mathbf{n}_2 \cdot \mathbf{d} = 0$, so $(a_1\mathbf{n}_1 + a_2\mathbf{n}_2) \cdot \mathbf{d} = 0$, which means $\mathbf{d} \cdot \mathbf{d} = 0$. But we can't have $\mathbf{d} \cdot \mathbf{d} = 0$ because $\mathbf{d}$ is nonzero! So, we conclude that $\mathbf{n}_1$ and $\mathbf{n}_2$ are linearly dependent.

□

Now that we have a thorough understanding of normal vectors, we turn to projections.

Problem 9.24: Let $\mathbf{v} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

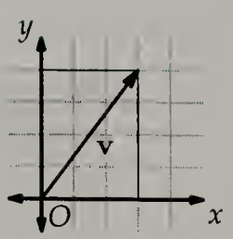
(a) Find the projections of $\mathbf{v}$ onto $\mathbf{i}$ and $\mathbf{j}$.

(b) Find the projection of $\mathbf{v}$ onto $2\mathbf{i}$.

(c) Find the projection of $\mathbf{v}$ onto $-776\mathbf{j}$.

Solution for Problem 9.24:

(a) To find the projections of $\mathbf{v}$ onto $\mathbf{i}$ and $\mathbf{j}$, we draw $\mathbf{v}$ from the origin to $(3, 4)$. When $\mathbf{i}$ is drawn from the origin, it lies along the x-axis. Therefore, we project $(3, 4)$ onto the x-axis to find the projection of $\mathbf{v}$ onto $\mathbf{i}$. As shown in the diagram at right, the projection of $(3, 4)$ onto the x-axis is $(3, 0)$, so the desired projection of $\mathbf{v}$ onto $\mathbf{i}$ is $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$. Similarly, the projection of $(3, 4)$ onto the y-axis is $(0, 4)$, so the projection of $\mathbf{v}$ onto $\mathbf{j}$ is $\begin{pmatrix} 0 \\ 4 \end{pmatrix}$.



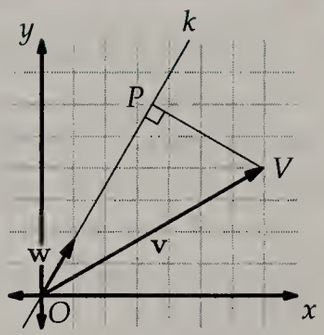
(b) Just as with $\mathbf{i}$, the vector $2\mathbf{i}$ lies along the x-axis when drawn from the origin. So, the projection of $\mathbf{v}$ onto $2\mathbf{i}$ is the same as the projection of $\mathbf{v}$ onto $\mathbf{i}$, which is $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$.

Concept: When $\mathbf{v}$ is projected onto vector $\mathbf{a}$, the magnitude of $\mathbf{a}$ is irrelevant; only its direction affects the projection.

(c) Only the direction of $-776\mathbf{j}$ affects the projection. When $-776\mathbf{j}$ is drawn from the origin, it lies along the y-axis, so the projection of $\mathbf{v}$ onto $-776\mathbf{j}$ is the same as the projection of $\mathbf{v}$ onto $\mathbf{j}$, which is $\begin{pmatrix} 0 \\ 4 \end{pmatrix}$.

Problem 9.25: Find the projection of $\mathbf{v} = \begin{pmatrix} 4\sqrt{3} \\ 4 \end{pmatrix}$ onto the vector $\mathbf{w} = \begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}$.

Solution for Problem 9.25: First, we draw $\mathbf{v}$ and $\mathbf{w}$ from the origin, and we draw the line k containing $\mathbf{w}$. Then, letting V be the head of $\mathbf{v}$, we project V onto k, which gives us point P as shown. Our goal is to find $\vec{P}$. We know the direction of $\vec{P}$, since it is the same as that of $\mathbf{w}$. So, we only have to find the magnitude of $\vec{P}$. We can do so with right triangle VOP, where O is the origin. In this triangle, we have $\cos\angle VOP = \frac{OP}{OV}$. But $OV = \|\mathbf{v}\| = 8$, so $\cos\angle VOP = \frac{OP}{8}$. If we can find $\cos\angle VOP$, then we can find OP. Since $\cos\angle VOP$ is the cosine of an angle between vectors $\mathbf{v}$ and $\mathbf{w}$, this sounds like a job for the dot product.



We have

$$\cos\angle VOP = \frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{v}\|\,\|\mathbf{w}\|} = \frac{(4\sqrt{3})(1)+(4)(\sqrt{3})}{(8)(2)} = \frac{\sqrt{3}}{2}.$$

Therefore, we have $OP = 8\cos\angle VOP = 4\sqrt{3}$. Since $\vec{P}$ is a vector in the same direction as $\mathbf{w}$ but with magnitude $4\sqrt{3}$, we find $\vec{P}$ by multiplying the unit vector in the direction of $\mathbf{w}$ by the magnitude of $\vec{P}$. The unit vector in the direction of $\mathbf{w}$ is $\frac{\mathbf{w}}{\|\mathbf{w}\|}$, so we have

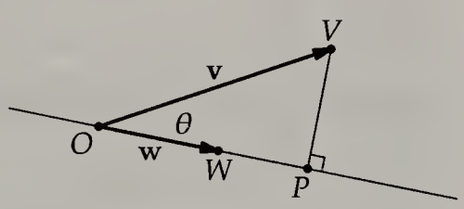
$$\vec{P} = 4\sqrt{3}\cdot\frac{\mathbf{w}}{\|\mathbf{w}\|} = 4\sqrt{3}\cdot\frac{\mathbf{w}}{2} = 2\sqrt{3}\mathbf{w} = \begin{pmatrix} 2\sqrt{3} \\ 6 \end{pmatrix}.$$

□

Problem 9.26: Find $\text{proj}_{\mathbf{w}}\mathbf{v}$ for any two vectors $\mathbf{v}$ and $\mathbf{w}$ in terms of $\mathbf{v}\cdot\mathbf{w}$, $\|\mathbf{w}\|$, and $\mathbf{w}$.

Solution for Problem 9.26: Let θ be the angle between $\mathbf{v}$ and $\mathbf{w}$. We'll handle the cases of acute and obtuse θ separately, and then address the special cases in which θ is 0, $\frac{\pi}{2}$, or π. In all cases, we draw $\mathbf{v}$ and $\mathbf{w}$ from a point O to points V and W, respectively, and let θ be the angle between these vectors. We then let P be the projection of V onto line $\overleftrightarrow{OW}$.

Case 1: θ is acute. We proceed as in Problem 9.25. We find $\vec{OP}$ by first finding $\|\vec{OP}\|$, and then scaling $\mathbf{w}$ by the appropriate factor. From right triangle VOP, we have $\|\vec{OP}\| = OV\cos\angle VOP = \|\mathbf{v}\|\cos\theta$. Since θ is the angle between $\mathbf{v}$ and $\mathbf{w}$, we have

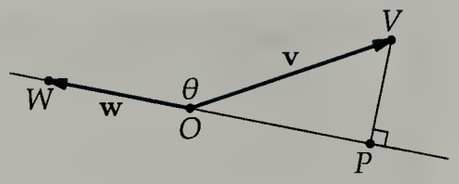


$$\|\vec{OP}\| = \|\mathbf{v}\|\cos\theta = \|\mathbf{v}\|\frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{v}\|\,\|\mathbf{w}\|} = \frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{w}\|}.$$

To find $\vec{OP}$, we scale the unit vector in the direction of $\mathbf{w}$ by $\|\vec{OP}\|$. From Problem 9.2, the unit vector in the direction of $\mathbf{w}$ is $\frac{\mathbf{w}}{\|\mathbf{w}\|}$, so

$$\text{proj}_{\mathbf{w}}\mathbf{v} = \vec{OP} = \|\vec{OP}\|\frac{\mathbf{w}}{\|\mathbf{w}\|} = \frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{w}\|^2}\mathbf{w}.$$

Case 2: θ is obtuse. We proceed as before, but now O is between W and P, as shown. So, $\left\|\vec{OP}\right\| = OV\cos\angle VOP = \|\mathbf{v}\|\cos(\pi-\theta) = -\|\mathbf{v}\|\cos\theta$. We then have



$$\left\|\vec{OP}\right\| = -\|\mathbf{v}\|\cos\theta = -\|\mathbf{v}\|\frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|} = -\frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{w}\|}.$$

Note that this result is positive, because $\mathbf{v}\cdot\mathbf{w}$ is negative when the angle between $\mathbf{v}$ and $\mathbf{w}$ is obtuse. To find $\vec{OP}$, we multiply the unit vector in the direction of $\mathbf{w}$ by $-\left\|\vec{OP}\right\|$, where the negative sign accounts for the fact that $\vec{OP}$ and $\mathbf{w}$ are in opposite directions. We therefore have

$$\text{proj}_{\mathbf{w}}\mathbf{v} = \vec{OP} = -\left\|\vec{OP}\right\|\frac{\mathbf{w}}{\|\mathbf{w}\|} = \frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{w}\|^2}\mathbf{w}.$$

This is the same expression as when θ is acute.

Case 3: $\theta = 0$ or $\theta = \pi$. In this case, we have $\mathbf{v} = k\mathbf{w}$ for some scalar k, and we should have $\text{proj}_{\mathbf{w}}\mathbf{v} = \mathbf{v}$, because $\mathbf{v}$ is already in the same (or opposite) direction as $\mathbf{w}$. The expression above for $\text{proj}_{\mathbf{w}}\mathbf{v}$ gives us

$$\frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{w}\|^2}\mathbf{w} = \frac{k\mathbf{w}\cdot(\mathbf{w})}{\|\mathbf{w}\|^2}\left(\frac{1}{k}\mathbf{v}\right) = \frac{k\|\mathbf{w}\|^2}{\|\mathbf{w}\|^2}\left(\frac{1}{k}\mathbf{v}\right) = \mathbf{v},$$

as expected.

Case 4: $\theta = \frac{\pi}{2}$. If $\mathbf{v}$ and $\mathbf{w}$ are orthogonal, then $\text{proj}_{\mathbf{w}}\mathbf{v} = \mathbf{0}$, since P will be the same point as O. The expression we derived for $\text{proj}_{\mathbf{w}}\mathbf{v}$ in the other cases delivers the correct answer in this case, too, because $\mathbf{v}\cdot\mathbf{w} = 0$ when $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

Since we've covered all the possible cases, we have shown that:

Important: For any vector $\mathbf{v}$ and any nonzero vector $\mathbf{w}$, we have

$$\text{proj}_{\mathbf{w}}\mathbf{v} = \frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{w}\|^2}\mathbf{w}.$$

□

We can test this formula for $\text{proj}_{\mathbf{w}}\mathbf{v}$ by confirming our result in Problem 9.25. Letting $\mathbf{v} = \begin{pmatrix}4\sqrt{3}\\4\end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix}1\\\sqrt{3}\end{pmatrix}$, we have

$$\text{proj}_{\mathbf{w}}\mathbf{v} = \frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{w}\|^2}\mathbf{w} = \frac{(4\sqrt{3})(1)+(4)(\sqrt{3})}{2^2}\begin{pmatrix}1\\\sqrt{3}\end{pmatrix} = \begin{pmatrix}2\sqrt{3}\\6\end{pmatrix},$$

which matches our answer from earlier.

Problem 9.27: Let $\mathbf{a} = 4\mathbf{i}+3\mathbf{j}$ and $\mathbf{b} = -\mathbf{i}+3\mathbf{j}$, and let $\mathbf{p} = \text{proj}_{\mathbf{b}}\mathbf{a}$.

(a) Find $\mathbf{p}$.

(b) Let $\mathbf{n} = \mathbf{a}-\mathbf{p}$. Find $\mathbf{b}\cdot\mathbf{n}$.

(c) Explain the geometric significance of your result in part (b).

Solution for Problem 9.27:

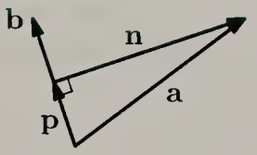
(a) We have $\text{proj}_{\mathbf{b}}\mathbf{a} = \dfrac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{b}\|^2}\mathbf{b} = \dfrac{(4)(-1)+(3)(3)}{\left(\sqrt{(-1)^2+3^2}\right)^2}\mathbf{b} = \dfrac{1}{2}\mathbf{b} = -\dfrac{1}{2}\mathbf{i}+\dfrac{3}{2}\mathbf{j}.$

(b) We have $\mathbf{n} = (4\mathbf{i} + 3\mathbf{j}) - \left(-\frac{1}{2}\mathbf{i} + \frac{3}{2}\mathbf{j}\right) = \frac{9}{2}\mathbf{i} + \frac{3}{2}\mathbf{j}$, and

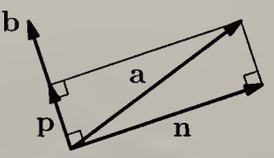
$$\mathbf{n} \cdot \mathbf{b} = \left(\frac{9}{2}\mathbf{i} + \frac{3}{2}\mathbf{j}\right) \cdot (-\mathbf{i} + 3\mathbf{j}) = \frac{9}{2} \cdot (-1) + \frac{3}{2} \cdot 3 = 0.$$

Therefore, $\mathbf{n}$ is orthogonal to $\mathbf{b}$.

(c) Writing $\mathbf{n} = \mathbf{a} - \mathbf{p}$ as $\mathbf{a} = \mathbf{p} + \mathbf{n}$ tells the story. Suppose we draw $\mathbf{a}$ and $\mathbf{b}$ from the origin. Because $\mathbf{p}$ is the projection of $\mathbf{a}$ onto $\mathbf{b}$, the vector from the head of $\mathbf{p}$ to the head of $\mathbf{a}$ must be normal to the line that contains $\mathbf{p}$. If we let this normal vector be $\mathbf{n}$, we must therefore have $\mathbf{a} = \mathbf{p} + \mathbf{n}$ and $\mathbf{n} \cdot \mathbf{p} = 0$. Since $\mathbf{b}$ and $\mathbf{p}$ are in the same or opposite directions, we must also have $\mathbf{n} \cdot \mathbf{b} = 0$.



Another way to look at this is that $\mathbf{n}$ is the projection of $\mathbf{a}$ onto a vector orthogonal to $\mathbf{b}$. This is depicted in the diagram at right. Therefore, we see that $\mathbf{a}$ is the sum of the projections of $\mathbf{a}$ onto $\mathbf{b}$ and onto a nonzero vector orthogonal to $\mathbf{b}$.



□

There's nothing special about the particular vectors $\mathbf{a}$ and $\mathbf{b}$ in Problem 9.27. We can make the same observations for any nonzero vectors $\mathbf{a}$ and $\mathbf{b}$.

You'll find a lot more use for projections and for expressing a vector as a sum of two orthogonal vectors when you study physics. We'll finish this section with an application to analytic geometry.

Problem 9.28: The distance between a point and a line is the length of the shortest possible segment from the point to the line.

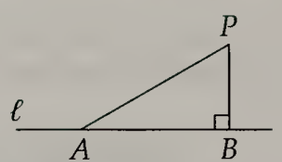
(a) Why does the distance between a point and a line equal the distance between the point and the projection of the point onto the line?

(b) Find the distance between $(4, -2)$ and the graph of $3x - 2y - 7 = 0$.

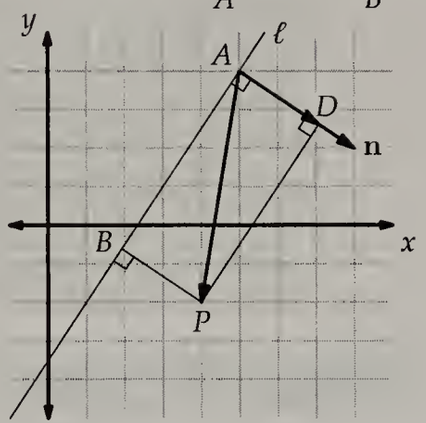
(c) Generalize your results in part (b) by finding an expression for the distance between the point (x_0, y_0) and the line $ax + by + c = 0$, where a, b, c, x_0, and y_0 are constants.

Solution for Problem 9.28:

(a) Let P be the point and ℓ be the line. Let B be the projection of P onto ℓ and let A be any other point on ℓ. The Pythagorean Theorem then gives us $PB^2 + AB^2 = PA^2$. Therefore, we have $PA > PB$. So, P is closer to B than to any other point on ℓ.



(b) Let the point be P and the line be ℓ. The shortest distance from P to ℓ is PB, where B is on ℓ such that $\overline{PB} \perp \ell$. Unfortunately, it's not clear from our diagram what the coordinates of B are. We could find those coordinates by finding the equation of the line through P perpendicular to ℓ, and then solving a system of equations. Then, we'd have to use the distance formula, and our diagram makes it obvious that we'll be dealing with a lot of fractions. That looks like a lot of work. Maybe we can find a better way.



Point B is the projection of P onto ℓ, and the vector from B to P is normal to ℓ. We know the direction of this vector! From our work in Problem 9.22, we know that $3\mathbf{i} - 2\mathbf{j}$ is normal to the graph of $3x - 2y - 7 = 0$. However, we don't know the magnitude of $\vec{BP}$ because we don't know exactly where B is. If only point B were a convenient point, then this problem would be a lot easier.

Let's take a look at a convenient point on ℓ, such as $(5, 4)$, which we'll call point A. This point is on ℓ because $(x, y) = (5, 4)$ satisfies $3x - 2y - 7 = 0$. How is this convenient point related to the desired point?

Triangle PAB is a right triangle. Therefore, not only is B the projection of P onto $\overleftrightarrow{AB}$, but B is also the projection of A onto $\overleftrightarrow{BP}$. Aha! We know the direction of $\overleftrightarrow{BP}$—the vector normal to ℓ is in this direction. So, we can find $\left\|\vec{BP}\right\|$ by projecting $\vec{AP}$ onto the normal vector!

In the diagram, we have drawn the vector $\mathbf{n}$ from A normal to ℓ, so $\mathbf{n} = 3\mathbf{i} - 2\mathbf{j}$. The projection of $\vec{AP}$ onto this vector, $\vec{AD}$, has magnitude equal to BP, since $BPDA$ is a rectangle. We have $\vec{AP} = \vec{P} - \vec{A} = (4\mathbf{i} - 2\mathbf{j}) - (5\mathbf{i} + 4\mathbf{j}) = -\mathbf{i} - 6\mathbf{j}$, and we're ready to find the desired distance:

$$BP = \left\| \frac{\vec{AP} \cdot \mathbf{n}}{\|\mathbf{n}\|^2}\mathbf{n} \right\| = \left| \frac{\vec{AP} \cdot \mathbf{n}}{\|\mathbf{n}\|^2} \right| \|\mathbf{n}\| = \left| \frac{(-1)(3) + (-6)(-2)}{\left(\sqrt{3^2 + (-2)^2}\right)^2} \right| \|\mathbf{n}\| = \frac{9}{13}\sqrt{13}.$$

(c) We use the previous part as a guide. Let P be (x_0, y_0) and let A be a point (x, y) on the graph of $ax + by + c = 0$. The vector normal to the graph is $\mathbf{n} = a\mathbf{i} + b\mathbf{j}$, and we seek the magnitude of the projection of $\vec{AP}$ onto $\mathbf{n}$. We have $\vec{AP} = (x_0 - x)\mathbf{i} + (y_0 - y)\mathbf{j}$, so the desired distance is

$$\left\| \frac{\vec{AP} \cdot \mathbf{n}}{\|\mathbf{n}\|^2}\mathbf{n} \right\| = \left| \frac{((x_0 - x)\mathbf{i} + (y_0 - y)\mathbf{j}) \cdot (a\mathbf{i} + b\mathbf{j})}{\|\mathbf{n}\|^2} \right| \|\mathbf{n}\| = \left| \frac{(x_0 - x)a + (y_0 - y)b}{\|\mathbf{n}\|^2} \right| \|\mathbf{n}\| = \frac{|ax_0 + by_0 - ax - by|}{\sqrt{a^2 + b^2}}.$$

Since (x, y) is on the graph of $ax + by + c = 0$, we have $-ax - by = c$, so our desired formula for the distance between (x_0, y_0) and $ax + by + c = 0$ is

$$\frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}.$$

□

Important: The distance between the point (x_0, y_0) and the graph of $ax + by + c = 0$ is

$$\frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}.$$

As one quick check, we can note that if (x_0, y_0) is on the graph of $ax + by + c = 0$, then $ax_0 + by_0 + c = 0$, and our formula for the distance between (x_0, y_0) and the line correctly gives 0.

WARNING!! Note that when using the formula

$$\frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}$$

to find the distance between (x_0, y_0) and $ax + by + c = 0$, all of the terms in the equation for the line are on the same side. If we have an equation such as $2x + 5y = 8$, then we do not have $a = 2$, $b = 5$ and $c = 8$. We move the 8 to the other side and have $2x + 5y - 8 = 0$, so $a = 2$, $b = 5$, and $c = -8$.

Exercises

9.4.1 Find $\text{proj}_{\mathbf{v}}\mathbf{u}$ for each of the following:

(a) $\mathbf{u} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$

(b) $\mathbf{u} = \begin{pmatrix} -3 \\ 1 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$

9.4.2 Suppose $\text{proj}_{\mathbf{v}}\mathbf{u} = 9\mathbf{i} - 3\mathbf{j}$.

(a) If $\mathbf{w} = 3\mathbf{v}$, then what is $\text{proj}_{\mathbf{w}}\mathbf{u}$?

(b) If $\mathbf{x} = 3\mathbf{u}$, then what is $\text{proj}_{\mathbf{v}}\mathbf{x}$?

9.4.3

(a) Find an equation of the form $ax + by + c = 0$ whose graph is a line through $(3, 4)$ such that $5\mathbf{i} - 3\mathbf{j}$ is normal to the line.

(b) Find a vector with magnitude 5 that is normal to the graph of $2x + 4y = 7$.

9.4.4 Show that if $\text{proj}_{\mathbf{x}}\mathbf{u} = \text{proj}_{\mathbf{x}}\mathbf{v}$ for all vectors $\mathbf{x}$, then $\mathbf{u} = \mathbf{v}$.

9.4.5 Let $\mathbf{v}$ be a nonzero vector.

(a) If $\text{proj}_{\mathbf{v}}\mathbf{w} = \text{proj}_{\mathbf{v}}(-\mathbf{w})$, then must $\mathbf{w} = \mathbf{0}$?

(b) If $\text{proj}_{\mathbf{v}}\mathbf{w} = \text{proj}_{(-\mathbf{v})}\mathbf{w}$, then must $\mathbf{w} = \mathbf{0}$?

9.4.6 Let $\mathbf{u}$ be the unit vector $x\mathbf{i} + y\mathbf{j}$, and let $\mathbf{u}^{\perp} = -y\mathbf{i} + x\mathbf{j}$.

(a) Find the angle between $\mathbf{u}$ and $\mathbf{u}^{\perp}$.

(b) Show that for any vector $\mathbf{a}$, we have $\mathbf{a} = \text{proj}_{\mathbf{u}}\mathbf{a} + \text{proj}_{\mathbf{u}^{\perp}}\mathbf{a}$. Explain the result both geometrically and algebraically.

(c) Show that $\|\mathbf{a}\|^2 = (\mathbf{a} \cdot \mathbf{u})^2 + (\mathbf{a} \cdot \mathbf{u}^{\perp})^2$.

9.4.7 Let θ be the angle between $\mathbf{u}$ and $\mathbf{v}$, where $\mathbf{u} \cdot \mathbf{v} \neq 0$. Find

$$\frac{\mathbf{u} \cdot \mathbf{v}}{(\text{proj}_{\mathbf{v}}\mathbf{u}) \cdot (\text{proj}_{\mathbf{u}}\mathbf{v})}$$

in terms of θ.

9.5 Summary

A **vector in two dimensions** is an ordered pair of numbers, which we can denote as $\begin{pmatrix} x \\ y \end{pmatrix}$. The numbers x and y are called **components** of the vector. We often refer to real numbers as **scalars** when speaking of vectors. The sum of of two vectors is given by $\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a+c \\ b+d \end{pmatrix}$, and the product of a scalar and a vector is given by $k\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ka \\ kb \end{pmatrix}$.

We denote the vector from the origin to a point C as $\vec{C}$. We refer to the vector from point A to point B as $\vec{AB}$, and we have $\vec{AB} = \vec{B} - \vec{A}$.

A vector with magnitude 1 is called a **unit vector**. The vector $\begin{pmatrix} x \\ y \end{pmatrix}$ can also be written as $x\mathbf{i} + y\mathbf{j}$, where $\mathbf{i}$ is the unit vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{j}$ is the unit vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. The **zero vector**, which is written as $\mathbf{0}$, is the vector in which both components are zero. We use the symbol $\mathbb{R}^2$ for the set of all vectors in two dimensions.

The **norm** of a vector $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ is $\|\mathbf{v}\| = \sqrt{x^2 + y^2}$. The norm of a vector is also called the **magnitude** or **length** of the vector, since it measures the length of the geometric representation of the vector.

We define the dot product $\mathbf{v} \cdot \mathbf{w}$ as $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = v_1 w_1 + v_2 w_2$. The dot product has the following properties:

- $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- $\mathbf{u} \cdot (a\mathbf{v}) = a(\mathbf{u} \cdot \mathbf{v})$
- $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
- $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$
- $\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \, \|\mathbf{w}\| \cos\theta$, where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$

We say that a function f with domain $\mathbb{R}^2$ is a **linear function** if $f(a\mathbf{v}) = af(\mathbf{v})$ for all scalars a and vectors $\mathbf{v}$, and $f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w})$ for all vectors $\mathbf{v}$ and $\mathbf{w}$.

Important: Each linear function g that maps vectors to scalars can be defined as $g(\mathbf{w}) = \mathbf{v} \cdot \mathbf{w}$ for a unique vector $\mathbf{v}$.

If $\mathbf{u}$ and $\mathbf{v}$ are vectors with $\mathbf{v} \neq \mathbf{0}$, then the graph of the parametric equation $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{u} + t\mathbf{v}$, is a **line**.

Definition: We say that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are **linearly dependent** if there are scalars $a_1, a_2, \ldots, a_n$, not all 0, such that

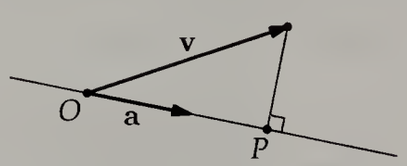
$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n = \mathbf{0}.$$

If the only scalars that satisfy this equation are $a_1 = a_2 = \cdots = a_n = 0$, then the vectors are **linearly independent**.

Important: Two vectors are linearly dependent if and only if one is a multiple of the other. Otherwise, they are linearly independent.

A vector is **normal** to a line if it is orthogonal to any vector in the direction of that line. The vector $\begin{pmatrix} a \\ b \end{pmatrix}$ is normal to the graph of $ax + by + c = 0$.

The **projection** of a point onto a line is the foot of the perpendicular from the point to the line. To find the projection of $\mathbf{v}$ onto a nonzero vector $\mathbf{a}$, we draw $\mathbf{v}$ from a point O, and then draw a line through O with direction vector $\mathbf{a}$. We then let P be the projection of the head of $\mathbf{v}$ onto this line. We call $\vec{OP}$ the projection of $\mathbf{v}$ onto $\mathbf{a}$. (Note that O does not have to be the origin in this definition!) We denote the projection of $\mathbf{v}$ onto $\mathbf{a}$ as $\text{proj}_{\mathbf{a}}\mathbf{v}$, so $\text{proj}_{\mathbf{a}}\mathbf{v} = \vec{OP}$.



Important: For any vector $\mathbf{v}$ and any nonzero vector $\mathbf{w}$, we have

$$\text{proj}_{\mathbf{w}}\mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{w}\|^2}\mathbf{w}.$$

Important: The distance between the point (x_0, y_0) and the graph of $ax + by + c = 0$ is

$$\frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}.$$

REVIEW PROBLEMS

9.29 Evaluate $\left\| \begin{pmatrix} -6 \\ 2 \end{pmatrix} \right\|$ and $\|15\mathbf{i} - 8\mathbf{j}\|$.

9.30 Find k such that $4\mathbf{i} - 5\mathbf{j}$ and $-3\mathbf{i} + k\mathbf{j}$ are linearly dependent.

9.31 Find a vector $\mathbf{v}$ such that $\|\mathbf{v}\| = 10$ and $\mathbf{v}$ is in the same direction as $4\mathbf{i} - 6\mathbf{j}$.

9.32 Let A, B, and C be points in the Cartesian plane. Find all possible values of $\vec{AB} + \vec{BC} + \vec{CA}$.

9.33 Points A, B, and C are in the Cartesian plane such that $\vec{C} = \frac{1}{2}(\vec{A} + \vec{B})$. How are A, B, and C related geometrically?

9.34 Show that $\|\mathbf{v}\| + \|\mathbf{w}\| \geq \|\mathbf{v} + \mathbf{w}\|$. What geometric result is this known as?

9.35 Evaluate each of the following

(a) $(3\mathbf{i} - 2\mathbf{j}) \cdot 8\mathbf{i}$ (b) $\begin{pmatrix} -2 \\ 5 \end{pmatrix} \cdot \begin{pmatrix} -4 \\ 3 \end{pmatrix}$

9.36 How are $\mathbf{u} \cdot \mathbf{v}$, $\|\mathbf{u}\|$, and $\|\mathbf{v}\|$ related if $\mathbf{u}$ and $\mathbf{v}$ have the same direction? What if they have opposite directions?

9.37

(a) Suppose $\mathbf{a}$ is a nonzero vector and $\mathbf{b}$ and $\mathbf{c}$ are vectors such that $\mathbf{a} \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{c} = 0$. Must $\mathbf{b}$ and $\mathbf{c}$ be linearly dependent?

(b) Suppose $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are nonzero vectors such that $\mathbf{a} \cdot \mathbf{b} = 0$ and $\mathbf{a} \cdot \mathbf{c} \neq 0$. Must $\mathbf{b}$ and $\mathbf{c}$ be linearly independent?

9.38 Let $\mathbf{u} = 3\mathbf{i} + 2\mathbf{j}$. Suppose each vector $\mathbf{v}$ such that $\mathbf{u} \cdot \mathbf{v} = 9$ is drawn from the origin in the Cartesian plane. How are the heads of all these vectors related?

9.39 Show that $\|\mathbf{v} + \mathbf{w}\| = \|\mathbf{v} - \mathbf{w}\|$ if and only if $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

9.40

(a) If two vectors have the same direction, but different magnitudes, then must one vector be a scalar multiple of the other?

(b) If one vector is a scalar multiple of the other, then must they have the same direction?

9.41 If $\mathbf{u} \cdot \mathbf{v}$ is nonzero, can we tell just from this dot product whether the angle between $\mathbf{u}$ and $\mathbf{v}$ is acute or obtuse?

9.42 Find $\text{proj}_{\mathbf{u}}\mathbf{v}$ for each of the following:

(a) $\mathbf{u} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 0 \\ -3 \end{pmatrix}$ (b) $\mathbf{u} = \begin{pmatrix} 5 \\ 7 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 2 \\ -4 \end{pmatrix}$

9.43 If $\text{proj}_{\mathbf{v}}\mathbf{u} = \mathbf{u}$, then must $\mathbf{u}$ and $\mathbf{v}$ be linearly dependent?

9.44 Find both an algebraic and a geometric proof that $\|\text{proj}_{\mathbf{b}}\mathbf{a}\| \leq \|\mathbf{a}\|$.

Challenge Problems

9.45 Prove that any three vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are linearly dependent.

9.46 How many vectors in two dimensions have one component equal to 2 and have magnitude 6? How are these vectors related geometrically?

9.47 Suppose that $\mathbf{u} \cdot \mathbf{v} = 2$ and $\mathbf{v} \cdot \mathbf{w} = -3$. What are the possible values of $\mathbf{u} \cdot \mathbf{w}$?

9.48 Suppose $\mathbf{a} = 3\mathbf{i} - \mathbf{j}$. **Hints:** 98

(a) Explain why more than one vector $\mathbf{b}$ satisfies $\text{proj}_{\mathbf{a}}\mathbf{b} = 6\mathbf{i} - 2\mathbf{j}$. Find two such vectors.

(b) Suppose we also know that $\|\mathbf{b}\| = 10$. Then, how many different vectors $\mathbf{b}$ satisfy $\text{proj}_{\mathbf{a}}\mathbf{b} = 6\mathbf{i} - 2\mathbf{j}$?

(c) Suppose that $\text{proj}_{\mathbf{j}}\mathbf{b} = -\mathbf{j}$. Then, how many different vectors $\mathbf{b}$ satisfy $\text{proj}_{\mathbf{a}}\mathbf{b} = 6\mathbf{i} - 2\mathbf{j}$?

9.49 Show that if $\mathbf{u}$ and $\mathbf{v}$ are linearly independent, then the values of $\text{proj}_{\mathbf{u}}\mathbf{r}$ and $\text{proj}_{\mathbf{v}}\mathbf{r}$ uniquely determine $\mathbf{r}$. In other words, if we are given values for $\text{proj}_{\mathbf{u}}\mathbf{r}$ and $\text{proj}_{\mathbf{v}}\mathbf{r}$, then there is always only one possible vector $\mathbf{r}$ that produces the given values of $\text{proj}_{\mathbf{u}}\mathbf{r}$ and $\text{proj}_{\mathbf{v}}\mathbf{r}$. **Hints:** 97

9.50 If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are nonzero vectors such that $\mathbf{u} = \text{proj}_{\mathbf{v}}\mathbf{u} + \text{proj}_{\mathbf{w}}\mathbf{u}$, then must we have $\mathbf{v} \cdot \mathbf{w} = 0$? **Hints:** 46

9.51★ Find the pair of equations whose graphs are the two lines that bisect the angles formed by the graphs of the lines $4x - 3y = 5$ and $x - 2y + 7 = 0$. **Hints:** 220

If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is.
– John von Neumann

CHAPTER 10
Matrices in Two Dimensions

10.1 What is a Matrix?

Consider the system of equations

$$\begin{aligned} 2x + 5y &= 1, \\ -3x - 4y &= -5. \end{aligned}$$

We can express this system with vectors by writing

$$\begin{pmatrix} 2x + 5y \\ -3x - 4y \end{pmatrix} = \begin{pmatrix} 1 \\ -5 \end{pmatrix}.$$

Each component on the left side is a linear combination of x and y. We have a special notation for expressing such a vector; we can express it as multiplication of $\begin{pmatrix} x \\ y \end{pmatrix}$ by a tool we call a **matrix**. Specifically, we can write

$$\begin{pmatrix} 2x + 5y \\ -3x - 4y \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ -3 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

where $\begin{pmatrix} 2 & 5 \\ -3 & -4 \end{pmatrix}$ is a matrix. (The plural of matrix is **matrices**.)

A matrix can be viewed as essentially a shorthand for keeping track of the coefficients of the variables in a system of linear equations. We say that the matrix above is a "2×2" matrix because it has 2 rows and 2 columns. In this section, we'll investigate some properties of 2×2 matrices, starting from the definition

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

Note that in this definition, the product of a matrix and a vector is a vector whose components are linear combinations of the components of the original vector.

More formally, a 2×2 matrix represents a function that has all vectors in two dimensions as its domain, and such that each output vector has components that are linear combinations of the input vector's components. The entries of the matrix describe how these linear combinations are formed, so the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ represents the function that outputs $\begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$ for each input $\begin{pmatrix} x \\ y \end{pmatrix}$. We express this function as the matrix multiplication

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

rather than with the usual function notation because the matrix tells us exactly how the output is determined for each input.

We will sometimes refer to a function that has vectors in two dimensions as inputs as having domain $\mathbb{R}^2$, and we'll say that a function "maps vectors to vectors" if it has vectors both as inputs and as outputs.

We typically use a bold capital letter to denote a matrix, and sometimes use the corresponding lowercase letter with subscripts to represent the entries in a matrix, like this:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

For each entry a_{ij}, the first number in the subscript indicates the row of the entry and the second number indicates the column, so a_{21} is the first entry of the second row. Also, just as we let $\mathbf{0}$ be the vector with both components equal to 0, we denote the matrix with all entries equal to 0 as $\mathbf{0}$, and call this the **zero matrix**.

Problems

Problem 10.1: Compute the following products:

(a) $\begin{pmatrix} 3 & -1 \\ 6 & 3 \end{pmatrix} \begin{pmatrix} -2 \\ 5 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & -3 \\ 5 & -1 \end{pmatrix} \begin{pmatrix} -5 \\ 1 \end{pmatrix}$

Problem 10.2: How are $\mathbf{Ai}$ and $\mathbf{Aj}$ related to $\mathbf{A}$?

Problem 10.3: We'd like to define multiplying a matrix by a scalar such that $(c\mathbf{A})\mathbf{v} = c(\mathbf{Av})$. How should we do so? That is, if c is a scalar and $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, then what matrix equals $c\mathbf{A}$?

Problem 10.4: Show that for any matrix $\mathbf{A}$ and any vectors $\mathbf{v}$ and $\mathbf{w}$, we have $\mathbf{A}(\mathbf{v} + \mathbf{w}) = \mathbf{Av} + \mathbf{Aw}$.

Problem 10.5: Let $\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 5 & -2 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 4 & -3 \\ -1 & 0 \end{pmatrix}$, and $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$.

(a) Find $\mathbf{Av}$ and $\mathbf{Bv}$ in terms of x and y.

(b) Find the matrix $\mathbf{C}$ such that $\mathbf{Cv} = \mathbf{Av} + \mathbf{Bv}$ for all $\mathbf{v}$.

(c) Let $\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$ and $\mathbf{Q} = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}$. Suppose we define $\mathbf{P} + \mathbf{Q}$ such that $(\mathbf{P} + \mathbf{Q})\mathbf{v} = \mathbf{Pv} + \mathbf{Qv}$ for all matrices $\mathbf{P}$ and $\mathbf{Q}$ and all vectors $\mathbf{v}$. What matrix equals $\mathbf{P} + \mathbf{Q}$?

We start with some practice multiplying vectors by matrices.

Problem 10.1: Compute the following products:

(a) $\begin{pmatrix} 3 & -1 \\ 6 & 3 \end{pmatrix}\begin{pmatrix} -2 \\ 5 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & -3 \\ 5 & -1 \end{pmatrix}\begin{pmatrix} -5 \\ 1 \end{pmatrix}$

Solution for Problem 10.1: We apply our definition of multiplication of a vector by a matrix:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}.$$

(a) $\begin{pmatrix} 3 & -1 \\ 6 & 3 \end{pmatrix}\begin{pmatrix} -2 \\ 5 \end{pmatrix} = \begin{pmatrix} (3)(-2)+(-1)(5) \\ (6)(-2)+(3)(5) \end{pmatrix} = \begin{pmatrix} -11 \\ 3 \end{pmatrix}.$

(b) $\begin{pmatrix} 0 & -3 \\ 5 & -1 \end{pmatrix}\begin{pmatrix} -5 \\ 1 \end{pmatrix} = \begin{pmatrix} (0)(-5)+(-3)(1) \\ (5)(-5)+(-1)(1) \end{pmatrix} = \begin{pmatrix} -3 \\ -26 \end{pmatrix}.$

□

WARNING!! Multiplication of a vector in two dimensions by a 2×2 matrix is only defined with the matrix on the left. We have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix},$$

but

$$\begin{pmatrix} x \\ y \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is *not* defined.

Problem 10.2: How are **Ai** and **Aj** related to **A**?

Solution for Problem 10.2: We have

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}.$$

Therefore, we see that:

Important: **Ai** is the first column of **A** and **Aj** is the second column of **A**.

You shouldn't have to memorize this—it's a straightforward application of our definition of the product of a matrix and a vector. However, sometimes this is a very useful way of thinking about matrices: the first column of **A** is **Ai** and the second column is **Aj**. □

Problem 10.3: We'd like to define multiplying a matrix by a scalar such that $(c\mathbf{A})\mathbf{v} = c(\mathbf{Av})$. How should we do so? In other words, if c is a scalar and $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, then what matrix equals $c\mathbf{A}$?

Solution for Problem 10.3: Just as with matrix addition, multiplying a matrix by a scalar works in the intuitive way, by multiplying each entry in the matrix by the scalar. We see this by noting that

$$c\left(\mathbf{A}\begin{pmatrix}x\\y\end{pmatrix}\right) = c\left(\begin{pmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\right) = c\begin{pmatrix}a_{11}x + a_{12}y\\a_{21}x + a_{22}y\end{pmatrix} = \begin{pmatrix}a_{11}cx + a_{12}cy\\a_{21}cx + a_{22}cy\end{pmatrix} = \begin{pmatrix}ca_{11} & ca_{12}\\ca_{21} & ca_{22}\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix},$$

so setting

$$c\mathbf{A} = \begin{pmatrix}ca_{11} & ca_{12}\\ca_{21} & ca_{22}\end{pmatrix}$$

gives us $(c\mathbf{A})\mathbf{v} = c(\mathbf{A}\mathbf{v})$, as desired. □

With this definition of multiplication of a matrix by a scalar, we can also show that $\mathbf{A}(c\mathbf{v}) = (c\mathbf{A})\mathbf{v} = c(\mathbf{A}\mathbf{v})$.

Important: For any scalar c, vector $\mathbf{v}$ and matrix $\mathbf{A}$, we have

$$(c\mathbf{A})\mathbf{v} = \mathbf{A}(c\mathbf{v}) = c(\mathbf{A}\mathbf{v}).$$

Problem 10.4: Show that for any matrix $\mathbf{A}$ and any vectors $\mathbf{v}$ and $\mathbf{w}$, we have $\mathbf{A}(\mathbf{v} + \mathbf{w}) = \mathbf{A}\mathbf{v} + \mathbf{A}\mathbf{w}$.

Solution for Problem 10.4: Writing out the details of $\mathbf{A}(\mathbf{v} + \mathbf{w})$, we have:

$$\begin{aligned}
\mathbf{A}(\mathbf{v} + \mathbf{w}) &= \begin{pmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{pmatrix}\left(\begin{pmatrix}v_1\\v_2\end{pmatrix} + \begin{pmatrix}w_1\\w_2\end{pmatrix}\right)\\
&= \begin{pmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}v_1 + w_1\\v_2 + w_2\end{pmatrix}\\
&= \begin{pmatrix}a_{11}v_1 + a_{11}w_1 + a_{12}v_2 + a_{12}w_2\\a_{21}v_1 + a_{21}w_1 + a_{22}v_2 + a_{22}w_2\end{pmatrix}\\
&= \begin{pmatrix}a_{11}v_1 + a_{12}v_2\\a_{21}v_1 + a_{22}v_2\end{pmatrix} + \begin{pmatrix}a_{11}w_1 + a_{12}w_2\\a_{21}w_1 + a_{22}w_2\end{pmatrix}\\
&= \begin{pmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}v_1\\v_2\end{pmatrix} + \begin{pmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}w_1\\w_2\end{pmatrix}\\
&= \mathbf{A}\mathbf{v} + \mathbf{A}\mathbf{w}.
\end{aligned}$$

Therefore, we have $\mathbf{A}(\mathbf{v} + \mathbf{w}) = \mathbf{A}\mathbf{v} + \mathbf{A}\mathbf{w}$. □

Problems 10.3 and 10.4 together tell us that the function that a matrix represents is a linear function. Letting f be one such function, this means that $f(a\mathbf{v}) = af(\mathbf{v})$ for all scalars a and vectors $\mathbf{v}$, and $f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v} + \mathbf{w})$ for all vectors $\mathbf{v}$ and $\mathbf{w}$. As we'll show in Problem 10.19, any linear function that maps a vector to a vector can be represented as multiplication by some matrix. Therefore, studying the basic properties of matrices is equivalent to studying linear functions that map vectors to vectors. Since linear functions are part of many fields of study, matrices are an essential mathematical tool with many applications.

Problem 10.5: Let $\mathbf{A} = \begin{pmatrix}1 & 3\\5 & -2\end{pmatrix}$, $\mathbf{B} = \begin{pmatrix}4 & -3\\-1 & 0\end{pmatrix}$, and $\mathbf{v} = \begin{pmatrix}x\\y\end{pmatrix}$.

(a) Find $\mathbf{A}\mathbf{v}$ and $\mathbf{B}\mathbf{v}$ in terms of x and y.

(b) Find the matrix $\mathbf{C}$ such that $\mathbf{C}\mathbf{v} = \mathbf{A}\mathbf{v} + \mathbf{B}\mathbf{v}$ for all $\mathbf{v}$.

(c) Let $\mathbf{P} = \begin{pmatrix}p_{11} & p_{12}\\p_{21} & p_{22}\end{pmatrix}$ and $\mathbf{Q} = \begin{pmatrix}q_{11} & q_{12}\\q_{21} & q_{22}\end{pmatrix}$. Suppose we define $\mathbf{P} + \mathbf{Q}$ such that $(\mathbf{P} + \mathbf{Q})\mathbf{v} = \mathbf{P}\mathbf{v} + \mathbf{Q}\mathbf{v}$ for all matrices $\mathbf{P}$ and $\mathbf{Q}$ and all vectors $\mathbf{v}$. What matrix equals $\mathbf{P} + \mathbf{Q}$?

Solution for Problem 10.5:

(a) We have

$$\mathbf{Av} = \begin{pmatrix} 1 & 3 \\ 5 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+3y \\ 5x-2y \end{pmatrix} \qquad \text{and} \qquad \mathbf{Bv} = \begin{pmatrix} 4 & -3 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4x-3y \\ -x \end{pmatrix}.$$

(b) Since $\mathbf{Av} + \mathbf{Bv} = \begin{pmatrix} x+3y \\ 5x-2y \end{pmatrix} + \begin{pmatrix} 4x-3y \\ -x \end{pmatrix} = \begin{pmatrix} 5x+0y \\ 4x-2y \end{pmatrix}$, the matrix $\mathbf{C}$ such that $\mathbf{Cv} = \mathbf{Av} + \mathbf{Bv}$ for all $\mathbf{v}$ is

$$\mathbf{C} = \begin{pmatrix} 5 & 0 \\ 4 & -2 \end{pmatrix}.$$

(c) We use our first two parts as a guide:

$$\begin{aligned} \mathbf{P}\begin{pmatrix} x \\ y \end{pmatrix} + \mathbf{Q}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} p_{11}x + p_{12}y \\ p_{21}x + p_{22}y \end{pmatrix} + \begin{pmatrix} q_{11}x + q_{12}y \\ q_{21}x + q_{22}y \end{pmatrix} \\ &= \begin{pmatrix} p_{11}x + p_{12}y + q_{11}x + q_{12}y \\ p_{21}x + p_{22}y + q_{21}x + q_{22}y \end{pmatrix} \\ &= \begin{pmatrix} (p_{11}+q_{11})x + (p_{12}+q_{12})y \\ (p_{21}+q_{21})x + (p_{22}+q_{22})y \end{pmatrix} \\ &= \begin{pmatrix} p_{11}+q_{11} & p_{12}+q_{12} \\ p_{21}+q_{21} & p_{22}+q_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \end{aligned}$$

So, we see that we should define $\mathbf{P} + \mathbf{Q}$ as

$$\mathbf{P} + \mathbf{Q} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} + \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} = \begin{pmatrix} p_{11}+q_{11} & p_{12}+q_{12} \\ p_{21}+q_{21} & p_{22}+q_{22} \end{pmatrix}.$$

□

We therefore see that adding matrices works just like adding vectors: we add two matrices by adding the corresponding entries in the two matrices.

Exercises

10.1.1 Find the following products:

(a) $\begin{pmatrix} 3 & -4 \\ 5 & -2 \end{pmatrix} \begin{pmatrix} 3 \\ -5 \end{pmatrix}$

(b) $\begin{pmatrix} -1 & 0 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} -1 \\ 3 \end{pmatrix}$

10.1.2 We define the difference of two matrices $\mathbf{A} - \mathbf{B}$ to be the matrix $\mathbf{D}$ such that $\mathbf{D} + \mathbf{B} = \mathbf{A}$. Show that if $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$, then

$$\mathbf{A} - \mathbf{B} = \begin{pmatrix} a_{11}-b_{11} & a_{12}-b_{12} \\ a_{21}-b_{21} & a_{22}-b_{22} \end{pmatrix}.$$

10.1.3 Let $\mathbf{A} = \begin{pmatrix} 6 & -4 \\ 3 & -2 \end{pmatrix}$.

(a) Find $\mathbf{A}\begin{pmatrix} 2 \\ -3 \end{pmatrix}$ and $\mathbf{A}\begin{pmatrix} 5 \\ -1 \end{pmatrix}$.

(b) Are $\begin{pmatrix} 2 \\ -3 \end{pmatrix}$ and $\begin{pmatrix} 5 \\ -1 \end{pmatrix}$ linearly dependent?

(c) Are your results in part (a) linearly dependent?

(d) If your answers to (b) and (c) are the same, then solve those parts again. If they are different, is it just a coincidence? If you choose any two vectors to replace $\begin{pmatrix} 2 \\ -3 \end{pmatrix}$ and $\begin{pmatrix} 5 \\ -1 \end{pmatrix}$ in part (a), will your answer to part (c) still be the same?

10.2 Multiplying Matrices

In the previous section, we found that matrix addition works in the way we might expect: each entry in the sum of two matrices is the sum of the corresponding entries in the two matrices. In this section, we learn how to find the product of two matrices, and find that the process isn't nearly as straightforward as adding two matrices.

Problems

Problem 10.6: Consider the system of equations

$$4x + 3y = 7,$$
$$2x - 9y = 5.$$

Suppose we wish to write the system in terms of the variables u and v, where $x = 2u - v$ and $y = 3u + v$.

(a) Write the original system in the form $\mathbf{A}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$.

(b) For what matrix $\mathbf{B}$ do we have $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{B}\begin{pmatrix} u \\ v \end{pmatrix}$?

(c) Use substitution to write the original system of equations in terms of u and v rather than x and y.

(d) Suppose we define the matrix product $\mathbf{AB}$ so that for any vector $\mathbf{w}$, we have $(\mathbf{AB})\mathbf{w} = \mathbf{A}(\mathbf{Bw})$. What is $\mathbf{AB}$?

Problem 10.7: Find a matrix that equals $\mathbf{AB}$, where

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Problem 10.8: Perform the following matrix multiplications:

(a) $\begin{pmatrix} -3 & 2 \\ -6 & -10 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & -5 \end{pmatrix}$

(b) $\begin{pmatrix} -1 & 0 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} -4 & 5 \\ 1 & -2 \end{pmatrix}$

Problem 10.9: Suppose $\mathbf{A}$ and $\mathbf{B}$ are 2×2 matrices.

(a) Must we have $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$?

(b) Must we have $\mathbf{AB} = \mathbf{BA}$?

Problem 10.10:

(a) Is matrix multiplication associative? In other words, is it always true that $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$?

(b) Is matrix multiplication distributive over addition? That is, is it always true that $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ and $(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$?

Problem 10.11:

(a) If $\mathbf{A}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}17\\0\end{pmatrix}$, then must $\mathbf{A} = \begin{pmatrix}3 & 7\\2 & -1\end{pmatrix}$?

(b) If $\mathbf{A}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}17\\0\end{pmatrix}$ and $\mathbf{A}\begin{pmatrix}-2\\-4\end{pmatrix} = \begin{pmatrix}-34\\0\end{pmatrix}$, then must $\mathbf{A} = \begin{pmatrix}3 & 7\\2 & -1\end{pmatrix}$?

(c) If $\mathbf{A}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}17\\0\end{pmatrix}$ and $\mathbf{A}\begin{pmatrix}2\\-1\end{pmatrix} = \begin{pmatrix}-1\\5\end{pmatrix}$, then must $\mathbf{A} = \begin{pmatrix}3 & 7\\2 & -1\end{pmatrix}$?

Problem 10.12: Suppose $\mathbf{v}$ and $\mathbf{w}$ are linearly independent and $\mathbf{A}$ and $\mathbf{B}$ are matrices such that $\mathbf{Av} = \mathbf{Bv}$ and $\mathbf{Aw} = \mathbf{Bw}$.

(a) Why must $\mathbf{Ai} = \mathbf{Bi}$?

(b) Show that $\mathbf{A} = \mathbf{B}$.

Problem 10.6: Consider the system of equations

$$4x + 3y = 7,$$
$$2x - 9y = 5.$$

Suppose we wish to write the system in terms of the variables u and v, where $x = 2u - v$ and $y = 3u + v$.

(a) Write the original system in the form $\mathbf{A}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}7\\5\end{pmatrix}$.

(b) For what matrix $\mathbf{B}$ do we have $\begin{pmatrix}x\\y\end{pmatrix} = \mathbf{B}\begin{pmatrix}u\\v\end{pmatrix}$?

(c) Use substitution to write the original system of equations in terms of u and v rather than x and y.

(d) Suppose we define the matrix product $\mathbf{AB}$ so that for any vector $\mathbf{w}$, we have $(\mathbf{AB})\mathbf{w} = \mathbf{A}(\mathbf{Bw})$. What is $\mathbf{AB}$?

Solution for Problem 10.6:

(a) We have $\begin{pmatrix}4 & 3\\2 & -9\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}7\\5\end{pmatrix}$, so $\mathbf{A} = \begin{pmatrix}4 & 3\\2 & -9\end{pmatrix}$.

(b) We have $\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}2 & -1\\3 & 1\end{pmatrix}\begin{pmatrix}u\\v\end{pmatrix}$, so $\mathbf{B} = \begin{pmatrix}2 & -1\\3 & 1\end{pmatrix}$.

(c) Substituting $x = 2u - v$ and $y = 3u + v$ into the original system of equations, and then doing some simplifying,

gives the system

$$17u - v = 7,$$
$$-23u - 11v = 5.$$

We can express this system as $\begin{pmatrix} 17 & -1 \\ -23 & -11 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$.

(d) Substituting $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$ from part (b) into $\begin{pmatrix} 4 & 3 \\ 2 & -9 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$ from part (a) gives

$$\begin{pmatrix} 4 & 3 \\ 2 & -9 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}.$$

Comparing this to part (c), it appears that we should define matrix multiplication such that

$$\begin{pmatrix} 4 & 3 \\ 2 & -9 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 17 & -1 \\ -23 & -11 \end{pmatrix}.$$

Notice that this is *not* what simple intuition might suggest.

□

WARNING!! ☢ We do *not* multiply two 2×2 matrices by multiplying corresponding entries. For example,

$$\begin{pmatrix} 4 & 3 \\ 2 & -9 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \neq \begin{pmatrix} 4 \cdot 2 & 3 \cdot (-1) \\ 2 \cdot 3 & -9 \cdot 1 \end{pmatrix}.$$

Let's see how matrix multiplication does work, given that we wish to define matrix multiplication so that $(\mathbf{AB})\,\mathbf{w} = \mathbf{A}\,(\mathbf{Bw})$ for all $\mathbf{w}$.

Problem 10.7: Find a matrix that equals $\mathbf{AB}$, where

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Solution for Problem 10.7: We let $\mathbf{w} = \begin{pmatrix} u \\ v \end{pmatrix}$, and we have

$$\begin{aligned}
\mathbf{A}\,(\mathbf{Bw}) &= \mathbf{A} \left(\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \right) \\
&= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11}u + b_{12}v \\ b_{21}u + b_{22}v \end{pmatrix} \\
&= \begin{pmatrix} a_{11}b_{11}u + a_{11}b_{12}v + a_{12}b_{21}u + a_{12}b_{22}v \\ a_{21}b_{11}u + a_{21}b_{12}v + a_{22}b_{21}u + a_{22}b_{22}v \end{pmatrix} \\
&= \begin{pmatrix} (a_{11}b_{11} + a_{12}b_{21})u + (a_{11}b_{12} + a_{12}b_{22})v \\ (a_{21}b_{11} + a_{22}b_{21})u + (a_{21}b_{12} + a_{22}b_{22})v \end{pmatrix} \\
&= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.
\end{aligned}$$

Therefore, we have $\mathbf{AB} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$. □

Important: We define the product of two 2×2 matrices as

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}.$$

We can remember how to multiply matrices by viewing each column of a matrix and each row of a matrix as a vector. The **column vectors** of $\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ are $\begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix}$ and $\begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix}$. Similarly, the two rows of a matrix **A** are the **row vectors** of **A**. We typically write row vectors horizontally, so the row vectors of $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ are $\begin{pmatrix} a_{11} & a_{12} \end{pmatrix}$ and $\begin{pmatrix} a_{21} & a_{22} \end{pmatrix}$.

With this in mind, we can remember how to perform matrix multiplication by noting that the entry of row i and column j of **AB** is the dot product of row i of **A** and column j of **B**. For example, the entry in row 1 and column 2 of **AB** is

$$\begin{pmatrix} a_{11} & a_{12} \end{pmatrix} \cdot \begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix} = a_{11}b_{12} + a_{12}b_{22}.$$

If we let the rows of **A** be the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$, and the columns of **B** be the vectors $\mathbf{b}_1$ and $\mathbf{b}_2$, we can therefore express the product **AB** as

$$\mathbf{AB} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{pmatrix} \begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 \end{pmatrix}.$$

We can also remember this matrix product by noting that each column of **AB** equals **A** times the corresponding column of **B**. That is, we can view the product as:

$$\mathbf{AB} = \mathbf{A} \begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{Ab}_1 & \mathbf{Ab}_2 \end{pmatrix}.$$

Problem 10.8: Perform the following matrix multiplications:

(a) $\begin{pmatrix} -3 & 2 \\ -6 & -10 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & -5 \end{pmatrix}$ (b) $\begin{pmatrix} -1 & 0 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} -4 & 5 \\ 1 & -2 \end{pmatrix}$

Solution for Problem 10.8:

(a)
$$\begin{pmatrix} -3 & 2 \\ -6 & -10 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & -5 \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} -3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix} & \begin{pmatrix} -3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -5 \end{pmatrix} \\ \begin{pmatrix} -6 & -10 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix} & \begin{pmatrix} -6 & -10 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -5 \end{pmatrix} \end{pmatrix}$$
$$= \begin{pmatrix} (-3)(0) + (2)(-1) & (-3)(1) + (2)(-5) \\ (-6)(0) + (-10)(-1) & (-6)(1) + (-10)(-5) \end{pmatrix}$$
$$= \begin{pmatrix} -2 & -13 \\ 10 & 44 \end{pmatrix}.$$

(b)
$$\begin{pmatrix} -1 & 0 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} -4 & 5 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} (-1)(-4) + (0)(1) & (-1)(5) + (0)(-2) \\ (4)(-4) + (3)(1) & (4)(5) + (3)(-2) \end{pmatrix} = \begin{pmatrix} 4 & -5 \\ -13 & 14 \end{pmatrix}.$$

□

Viewing matrices as representing functions, the matrix product **AB** is a composition of functions. The product $\mathbf{A}\left(\mathbf{B}\begin{pmatrix} x \\ y \end{pmatrix}\right)$ tells us to apply the function that **B** represents to $\begin{pmatrix} x \\ y \end{pmatrix}$, and then apply the function that **A** represents

to the result of $\mathbf{B}\begin{pmatrix} x \\ y \end{pmatrix}$. As we have just seen, we can represent the product **AB** as a single matrix, and this matrix represents the composition of the functions **A** and **B** represent. So, the product $(\mathbf{AB})\begin{pmatrix} x \\ y \end{pmatrix}$ tells us to apply the result of this function composition to $\begin{pmatrix} x \\ y \end{pmatrix}$.

Problem 10.9: Suppose **A** and **B** are 2×2 matrices.

(a) Must we have $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$?

(b) Must we have $\mathbf{AB} = \mathbf{BA}$?

Solution for Problem 10.9:

(a) Yes. We have

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}$$
$$= \begin{pmatrix} b_{11} + a_{11} & b_{12} + a_{12} \\ b_{21} + a_{21} & b_{22} + a_{22} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \mathbf{B} + \mathbf{A}.$$

(b) The non-intuitive definition of matrix multiplication might make us suspect that **AB** is not necessarily equal to **BA**. We can test it by looking at one of the examples in the previous problem. For example, we found that

$$\begin{pmatrix} -1 & 0 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} -4 & 5 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} 4 & -5 \\ -13 & 14 \end{pmatrix}.$$

Reversing the order of the matrices in the product, we have

$$\begin{pmatrix} -4 & 5 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} (-4)(-1) + (5)(4) & (-4)(0) + (5)(3) \\ (1)(-1) + (-2)(4) & (1)(0) + (-2)(3) \end{pmatrix} = \begin{pmatrix} 24 & 15 \\ -9 & -6 \end{pmatrix}.$$

So, on our first try we found a pair of matrices **A** and **B** such that **AB** and **BA** are not equal.

□

WARNING!! Matrix multiplication is not commutative; the products **AB** and **BA** are not necessarily equal.

Sidenote: When we think of a matrix as representing a function, it's not surprising that matrix multiplication is not commutative. The product $\mathbf{AB}\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{A}\left(\mathbf{B}\begin{pmatrix} x \\ y \end{pmatrix}\right)$ tells us to apply the function that **B** represents to $\begin{pmatrix} x \\ y \end{pmatrix}$, and then apply the function that **A** represents to the result of $\mathbf{B}\begin{pmatrix} x \\ y \end{pmatrix}$. The product $\mathbf{BA}\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{B}\left(\mathbf{A}\begin{pmatrix} x \\ y \end{pmatrix}\right)$ has us apply these functions in the reverse order. But we saw on page 17 that function composition is not commutative; if f and g are functions, then $f \circ g$ is not necessarily the same as $g \circ f$.

Problem 10.10:

(a) Is matrix multiplication associative? In other words, is it always true that $\mathbf{A(BC) = (AB)C}$?

(b) Is matrix multiplication distributive over addition? That is, is it always true that $\mathbf{A(B + C) = AB + AC}$ and $\mathbf{(B + C)A = BA + CA}$?

Solution for Problem 10.10:

(a) We have

$$\begin{aligned}\mathbf{A(BC)} &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11}c_{11} + b_{12}c_{21} & b_{11}c_{12} + b_{12}c_{22} \\ b_{21}c_{11} + b_{22}c_{21} & b_{21}c_{12} + b_{22}c_{22} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{11}c_{11} + a_{11}b_{12}c_{21} + a_{12}b_{21}c_{11} + a_{12}b_{22}c_{21} & a_{11}b_{11}c_{12} + a_{11}b_{12}c_{22} + a_{12}b_{21}c_{12} + a_{12}b_{22}c_{22} \\ a_{21}b_{11}c_{11} + a_{21}b_{12}c_{21} + a_{22}b_{21}c_{11} + a_{22}b_{22}c_{21} & a_{21}b_{11}c_{12} + a_{21}b_{12}c_{22} + a_{22}b_{21}c_{12} + a_{22}b_{22}c_{22} \end{pmatrix}\end{aligned}$$

and

$$\begin{aligned}\mathbf{(AB)C} &= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{11}c_{11} + a_{12}b_{21}c_{11} + a_{11}b_{12}c_{21} + a_{12}b_{22}c_{21} & a_{11}b_{11}c_{12} + a_{12}b_{21}c_{12} + a_{11}b_{12}c_{22} + a_{12}b_{22}c_{22} \\ a_{21}b_{11}c_{11} + a_{22}b_{21}c_{11} + a_{21}b_{12}c_{21} + a_{22}b_{22}c_{21} & a_{21}b_{11}c_{12} + a_{22}b_{21}c_{12} + a_{21}b_{12}c_{22} + a_{22}b_{22}c_{22} \end{pmatrix}.\end{aligned}$$

With a little tedious inspection, we can see that these expressions for $\mathbf{A(BC)}$ and $\mathbf{(AB)C}$ are indeed the same. Therefore, we can simply write $\mathbf{ABC}$ rather than writing $\mathbf{A(BC)}$ or $\mathbf{(AB)C}$.

Our proof that $\mathbf{A(BC) = (AB)C}$ is not particularly illuminating. We can find a more intuitively satisfying explanation by starting with systems of equations:

$$\begin{aligned} t_1 &= a_{11}u_1 + a_{12}u_2, & u_1 &= b_{11}v_1 + b_{12}v_2, & v_1 &= c_{11}w_1 + c_{12}w_2, \\ t_2 &= a_{21}u_1 + a_{22}u_2, & u_2 &= b_{21}v_1 + b_{22}v_2, & v_2 &= c_{21}w_1 + c_{22}w_2. \end{aligned}$$

We can view these systems as $\mathbf{t = Au}$, $\mathbf{u = Bv}$, and $\mathbf{v = Cw}$. Through a pair of substitutions, we can combine these systems to express t_1 and t_2 in terms of w_1 and w_2. No matter which order we perform these substitutions, we'll get the same system of equations in the end. If we substitute the expressions for u_1 and u_2 from the second system into the first, and then substitute the expressions for v_1 and v_2 into the result, we get $\mathbf{t = (AB)(Cw)}$. If we instead first substitute the expressions for v_1 and v_2 from the the third system into the second, and then substitute the result into the first, we get $\mathbf{t = A(BCw)}$. Either way, we get the same system.

(b) This one's a good deal easier. We have

$$\begin{aligned}\mathbf{A(B + C)} &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} + c_{11} & b_{12} + c_{12} \\ b_{21} + c_{21} & b_{22} + c_{22} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{11} + a_{11}c_{11} + a_{12}b_{21} + a_{12}c_{21} & a_{11}b_{12} + a_{11}c_{12} + a_{12}b_{22} + a_{12}c_{22} \\ a_{21}b_{11} + a_{21}c_{11} + a_{22}b_{21} + a_{22}c_{21} & a_{21}b_{12} + a_{21}c_{12} + a_{22}b_{22} + a_{22}c_{22} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} + \begin{pmatrix} a_{11}c_{11} + a_{12}c_{21} & a_{11}c_{12} + a_{12}c_{22} \\ a_{21}c_{11} + a_{22}c_{21} & a_{21}c_{12} + a_{22}c_{22} \end{pmatrix} \\ &= \mathbf{AB + AC},\end{aligned}$$

and similar computations show that $\mathbf{(B + C)A = BA + CA}$. (Note that we must confirm $\mathbf{A(B + C) = AB + AC}$ and $\mathbf{(B + C)A = BA + CA}$ separately because matrix multiplication is not commutative.)

□

Important: Matrix multiplication is associative and distributive. So, for all 2×2 matrices **A**, **B**, and **C**, we have

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C},$$
$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC},$$
$$(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}.$$

Sidenote: Again, we can use the fact that a matrix represents a function to understand why matrix multiplication is associative. For any three functions f, g, and h, we have

$$f \circ (g \circ h) = (f \circ g) \circ h.$$

If matrices **A**, **B**, and **C** represent functions f, g, and h, respectively, then we have

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}.$$

Problem 10.11:

(a) If $\mathbf{A}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}17\\0\end{pmatrix}$, then must $\mathbf{A} = \begin{pmatrix}3 & 7\\2 & -1\end{pmatrix}$?

(b) If $\mathbf{A}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}17\\0\end{pmatrix}$ and $\mathbf{A}\begin{pmatrix}-2\\-4\end{pmatrix} = \begin{pmatrix}-34\\0\end{pmatrix}$, then must $\mathbf{A} = \begin{pmatrix}3 & 7\\2 & -1\end{pmatrix}$?

(c) If $\mathbf{A}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}17\\0\end{pmatrix}$ and $\mathbf{A}\begin{pmatrix}2\\-1\end{pmatrix} = \begin{pmatrix}-1\\5\end{pmatrix}$, then must $\mathbf{A} = \begin{pmatrix}3 & 7\\2 & -1\end{pmatrix}$?

Solution for Problem 10.11: Let $\mathbf{A} = \begin{pmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{pmatrix}$, so $\mathbf{A}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}a_{11} + 2a_{12}\\a_{21} + 2a_{22}\end{pmatrix}$.

(a) We must have $a_{11} + 2a_{12} = 17$ and $a_{21} + 2a_{22} = 0$. We have four variables among these two linear equations. Choosing $a_{12} = 7$ and $a_{22} = -1$, and then solving for a_{11} and a_{21}, gives us $\mathbf{A} = \begin{pmatrix}3 & 7\\2 & -1\end{pmatrix}$. But there are other possibilities for **A**, since these aren't the only possible values we can choose for a_{12} and a_{22}. The solution $\mathbf{A} = \begin{pmatrix}17 & 0\\0 & 0\end{pmatrix}$ is one particularly obvious possibility for **A**. Choosing $a_{12} = 1$ and $a_{22} = 2$, and then solving for a_{11} and a_{21}, gives us $\mathbf{A} = \begin{pmatrix}15 & 1\\-4 & 2\end{pmatrix}$ as another possibility. For any choices of a_{12} and a_{22}, we can solve for a_{11} and a_{21} to produce a matrix that satisfies $\mathbf{A}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}17\\0\end{pmatrix}$, so there are infinitely many possibilities for **A**.

(b) Now, we have four equations for our four variables. As in part (a), we have

$$a_{11} + 2a_{12} = 17,$$
$$a_{21} + 2a_{22} = 0.$$

From $\mathbf{A}\begin{pmatrix}-2\\-4\end{pmatrix} = \begin{pmatrix}-34\\0\end{pmatrix}$, we also have

$$-2a_{11} - 4a_{12} = -34,$$
$$-2a_{21} - 4a_{22} = 0.$$

Adding twice $a_{11} + 2a_{12} = 17$ to $-2a_{11} - 4a_{12} = -34$ gives us $0 = 0$. Uh-oh. The same happens when we add twice $a_{21} + 2a_{22} = 0$ to $-2a_{21} - 4a_{22} = 0$. The second system of equations above is just -2 times the first

system! Since the new pair of equations is just the same as the pair of equations we already had, all the solutions in part (a) are also solutions here. Once again, there is not just one possible matrix **A**.

(c) Again, we must have $a_{11} + 2a_{12} = 17$ and $a_{21} + 2a_{22} = 0$. From $\mathbf{A}\begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 5 \end{pmatrix}$, we have $2a_{11} - a_{12} = -1$ and $2a_{21} - a_{22} = 5$. The equations $2a_{11} - a_{12} = -1$ and $a_{11} + 2a_{12} = 17$ give us $a_{11} = 3$ and $a_{12} = 7$, and the other two equations give us $a_{21} = 2$ and $a_{22} = -1$. So, in this part, we must indeed have $\mathbf{A} = \begin{pmatrix} 3 & 7 \\ 2 & -1 \end{pmatrix}$.

□

What made part (c) different from the other two parts? It isn't surprising that knowing what $\mathbf{Av}$ equals for one vector $\mathbf{v}$ isn't enough to figure out what **A** is, but why can we figure out what **A** is in part (c), but not in part (b)? Looking more closely at (b), we see that the two vectors multiplied by **A** are $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} -2 \\ -4 \end{pmatrix}$, which are not linearly independent. But the two vectors multiplied by **A** in part (c) are linearly independent. Maybe that's the key? Let's see.

Problem 10.12: Suppose $\mathbf{v}$ and $\mathbf{w}$ are linearly independent and **A** and **B** are matrices such that $\mathbf{Av} = \mathbf{Bv}$ and $\mathbf{Aw} = \mathbf{Bw}$.

(a) Why must $\mathbf{Ai} = \mathbf{Bi}$?

(b) Show that $\mathbf{A} = \mathbf{B}$.

Solution for Problem 10.12:

(a) In Problem 9.21, we showed that if $\mathbf{v}$ and $\mathbf{w}$ are linearly independent, then every point (x, y) is in the graph of the parametric equation $\begin{pmatrix} x \\ y \end{pmatrix} = p\mathbf{v} + q\mathbf{w}$, where p and q are parameters. Therefore, for any vector $\mathbf{u}$, there are scalars p and q such that $p\mathbf{v} + q\mathbf{w} = \mathbf{u}$. Specifically, there are scalars r and s such that $r\mathbf{v} + s\mathbf{w} = \mathbf{i}$. Adding r times $\mathbf{Av} = \mathbf{Bv}$ to s times $\mathbf{Aw} = \mathbf{Bw}$ gives $r\mathbf{Av} + s\mathbf{Aw} = r\mathbf{Bv} + s\mathbf{Bw}$, which we can manipulate as follows:

$$\begin{aligned} & r\mathbf{Av} + s\mathbf{Aw} = r\mathbf{Bv} + s\mathbf{Bw} \\ \Rightarrow \quad & \mathbf{A}(r\mathbf{v}) + \mathbf{A}(s\mathbf{w}) = \mathbf{B}(r\mathbf{v}) + \mathbf{B}(s\mathbf{w}) \\ \Rightarrow \quad & \mathbf{A}(r\mathbf{v} + s\mathbf{w}) = \mathbf{B}(r\mathbf{v} + s\mathbf{w}). \end{aligned}$$

We also know that $r\mathbf{v} + s\mathbf{w} = \mathbf{i}$, so we have $\mathbf{Ai} = \mathbf{Bi}$.

(b) In Problem 10.2, we saw that the first column of **A** is $\mathbf{Ai}$. Therefore, part (a) tells us that the first column of **A** is the same as the first column of **B**. By a similar argument to part (a), we can show that $\mathbf{Aj} = \mathbf{Bj}$, so the second column of **A** is the same as the second column of **B**. Since **A** and **B** have the same first column and the same second column, we have $\mathbf{A} = \mathbf{B}$.

□

Important: If $\mathbf{v}$ and $\mathbf{w}$ are linearly independent and **A** and **B** are matrices such that $\mathbf{Av} = \mathbf{Bv}$ and $\mathbf{Aw} = \mathbf{Bw}$, then $\mathbf{A} = \mathbf{B}$.

Exercises

10.2.1 Compute the following products:

(a) $\begin{pmatrix} 3 & 4 \\ -2 & -6 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ -5 & -7 \end{pmatrix}$

(b) $\begin{pmatrix} -4 & -8 \\ 0 & 9 \end{pmatrix}\begin{pmatrix} -\frac{1}{2} & 2 \\ -3 & 3 \end{pmatrix}$

10.2.2 A 2×2 matrix is **diagonal** if it is of the form $\begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$. Show that the product of two diagonal matrices is also diagonal.

10.2.3 Is there a matrix **F** such that for any matrix **A**, the matrix **FA** is the matrix formed by swapping the rows of **A**? If such a matrix exists, what is **AF** for that matrix?

10.2.4 If $\mathbf{AB} = \mathbf{0}$, then must either **A** or **B** be **0**?

10.2.5 Let $\mathbf{T}\begin{pmatrix} 4 \\ -3 \end{pmatrix} = \begin{pmatrix} -3 \\ 4 \end{pmatrix}$ and $\mathbf{T}\begin{pmatrix} -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \end{pmatrix}$. What is $\mathbf{T}\begin{pmatrix} 4 & -3 \\ -3 & 1 \end{pmatrix}$?

10.2.6 In Problem 10.11, we showed that there is more than one matrix **A** for which we have $\mathbf{A}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 17 \\ 0 \end{pmatrix}$ and $\mathbf{A}\begin{pmatrix} -2 \\ -4 \end{pmatrix} = \begin{pmatrix} -34 \\ 0 \end{pmatrix}$. Our first step in doing so was letting $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and building systems of equations by multiplying out the left sides of the given equations. Did we have to do that in order to see that the two given equations are really the same equation?

10.3 Matrices as Transformations

Problems

Problem 10.13:

(a) Find a matrix **I** such that multiplying any vector by **I** leaves the vector unchanged. In other words, find **I** such that $\mathbf{Iv} = \mathbf{v}$ for all vectors **v**. Is this matrix unique?

(b) Do we have $\mathbf{IA} = \mathbf{A}$ for any matrix **A**?

Problem 10.14: Find the matrix **A** such that $\mathbf{Av} = k\mathbf{v}$ for all vectors **v**, where k is a scalar.

Problem 10.15: In this problem, we show that there is a matrix **N** such that **Nv** is a 90° counterclockwise rotation of **v** for all **v**.

(a) Let $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$. Find the image of the point (x, y) under a 90° counterclockwise rotation. Use your result to find the desired matrix **N**.

(b) Use part (a) to show that a 90° counterclockwise rotation is a linear transformation. In other words, show that the image of a 90° counterclockwise rotation of $c\mathbf{v}$ is c times the 90° counterclockwise rotation of **v**, and the image of a 90° counterclockwise rotation of $\mathbf{v} + \mathbf{w}$ is the sum of the 90° counterclockwise rotations of **v** and **w**.

Problem 10.16: In this problem, we find a matrix **T** such that the vector **Tv** is a counterclockwise rotation of the vector **v** by an angle ϕ.

(a) Let $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$. Suppose the point P is the image upon rotating the point (x, y) an angle ϕ counterclockwise about the origin. What are the coordinates of P?

(b) Find the desired matrix **T**. Test your matrix by finding **T** for $\phi = 0°$ and $\phi = 90°$.

Problem 10.17: A sequence $(a_1, b_1), (a_2, b_2), (a_3, b_3), \ldots$ of points in the coordinate plane satisfies

$$(a_{n+1}, b_{n+1}) = (\sqrt{3}a_n - b_n, \sqrt{3}b_n + a_n)$$

for $n = 1, 2, 3, \ldots$. Suppose that $(a_{100}, b_{100}) = (2, 4)$. What is $a_1 + b_1$? *(Source: AMC 12)*

Problem 10.18: Let $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ and z be the complex number $x + yi$.

(a) Find the matrix $\mathbf{C}$ such that $\mathbf{Cv} = \begin{pmatrix} \text{Re}(zi) \\ \text{Im}(zi) \end{pmatrix}$. In other words, find the matrix $\mathbf{C}$ such that multiplying $\mathbf{v}$ by $\mathbf{C}$ corresponds to multiplying z by i.

(b) Find the matrix $\mathbf{D}$ such that multiplying $\mathbf{v}$ by $\mathbf{D}$ corresponds to multiplying z by $a + bi$. (Your answer should be a matrix whose entries are in terms of a and b.)

Problem 10.19: As a reminder, we say that a function f with domain $\mathbb{R}^2$ is a linear function if

(i) $f(a\mathbf{v}) = af(\mathbf{v})$ for all scalars a and vectors $\mathbf{v}$, and

(ii) $f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w})$ for all vectors $\mathbf{v}$ and $\mathbf{w}$.

Suppose $g(\mathbf{w})$ is a linear function that maps vectors to vectors.

(a) Explain why $g(x\mathbf{i} + y\mathbf{j}) = xg(\mathbf{i}) + yg(\mathbf{j})$.

(b) Why does part (a) tell us that there exists some matrix $\mathbf{A}$ such that $g(\mathbf{w}) = \mathbf{Aw}$ for all $\mathbf{w}$?

Problem 10.13:

(a) Find the matrix $\mathbf{I}$ such that multiplying any vector by $\mathbf{I}$ leaves the vector unchanged. In other words, find $\mathbf{I}$ such that $\mathbf{Iv} = \mathbf{v}$ for all vectors $\mathbf{v}$. Is this matrix unique?

(b) Do we have $\mathbf{IA} = \mathbf{AI} = \mathbf{A}$ for any matrix $\mathbf{A}$?

Solution for Problem 10.13:

(a) From Problem 10.2, we know that the first column of $\mathbf{I}$ is $\mathbf{Ii}$. Since $\mathbf{Iv} = \mathbf{v}$ for all $\mathbf{v}$, we have $\mathbf{Ii} = \mathbf{i}$, so the first column of $\mathbf{I}$ is $\mathbf{i}$. Similarly, we have $\mathbf{Ij} = \mathbf{j}$, so the second column of $\mathbf{I}$ must be $\mathbf{j}$. So, if there is a matrix $\mathbf{I}$ such that $\mathbf{Iv} = \mathbf{v}$, then it must be

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We aren't quite finished! We have only shown that if there is a matrix such that $\mathbf{Iv} = \mathbf{v}$, then it must be $\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. We have not shown that this matrix satisfies $\mathbf{Iv} = \mathbf{v}$ for all $\mathbf{v}$; we've only shown that $\mathbf{Iv} = \mathbf{v}$ when $\mathbf{v}$ is $\mathbf{i}$ or $\mathbf{j}$. Fortunately, we can patch this hole very quickly:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix},$$

so $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{v} = \mathbf{v}$ for all $\mathbf{v}$.

WARNING!! Make sure you see why the second part of our solution is necessary. Our first step of considering **Ii** and **Ij** only tells us that what the columns of **I** must be *if a matrix with the desired property exists.* We need the last part of the solution to show that the matrix thus constructed does indeed satisfy **Iv** = **v** for all **v**.

(b) First, we show that **IA** = **A**. The first column of **IA** is the product of **I** and the first column of **A**, which is just the first column of **A**. Similarly, the second column of **IA** must be the second column of **A**, so **IA** = **A**. We also could have written out both matrices to confirm:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

For yet another solution, note that part (a) gives us **I**(**Av**) = **Av** for any matrix **A** and vector **v**. We also have **I**(**Av**) = (**IA**)**v**. Therefore, we must have (**IA**)**v** = **Av** for all vectors **v**, so **IA** = **A**.

To show that **AI** = **A**, we can again write out both matrices:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

We could also again use part (a). We have (**AI**)**v** = **A**(**Iv**) = **Av** for all vectors **v**, so we have **AI** = **A**.

□

Important: The matrix

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is called the **identity matrix**. Multiplying a vector **v** by **I** leaves the vector unchanged; in other words, **Iv** = **v** for all vectors **v**. We also have **IA** = **AI** = **A** for any 2×2 matrix **A**.

Problem 10.14: Find the matrix **A** such that **Av** = k**v** for all vectors **v**, where k is a scalar.

Solution for Problem 10.14: Solution 1: Equate corresponding components. Performing the multiplications on both sides of

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = k \begin{pmatrix} x \\ y \end{pmatrix}$$

gives

$$\begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} = \begin{pmatrix} kx \\ ky \end{pmatrix}.$$

From $ax + by = kx$, we have $a = k$ and $b = 0$, and from $cx + dy = ky$, we have $c = 0$ and $d = k$. Checking, we see that

$$\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} kx \\ ky \end{pmatrix} = k \begin{pmatrix} x \\ y \end{pmatrix}$$

for all x and y, as desired.

Solution 2: Apply the result of Problem 10.2. We must have **Av** = k**v** for all **v**, so we must have **Ai** = k**i** and **Aj** = k**j**. Therefore, if there is a matrix **A** such that **Av** = k**v** for all **v**, then the first column of **A** must be k**i** and the second must be k**j**. We can quickly confirm that $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = k \begin{pmatrix} x \\ y \end{pmatrix}$, so $\mathbf{A} = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$ is the desired matrix. □

So far, we have discovered a matrix that does nothing (the identity) and a family of matrices that scale vectors. Now, we turn to matrices that rotate vectors.

Problem 10.15: In this problem, we show that there is a matrix $\mathbf{N}$ such that $\mathbf{Nv}$ is a 90° counterclockwise rotation of $\mathbf{v}$ for all $\mathbf{v}$.

(a) Let $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$. Find the image of the point (x, y) under a 90° counterclockwise rotation. Use your result to find the desired matrix $\mathbf{N}$.

(b) Use part (a) to show that a 90° counterclockwise rotation is a linear transformation. In other words, show that the image of a 90° counterclockwise rotation of $c\mathbf{v}$ is c times the 90° counterclockwise rotation of $\mathbf{v}$, and the image of a 90° counterclockwise rotation of $\mathbf{v} + \mathbf{w}$ is the sum of the 90° counterclockwise rotations of $\mathbf{v}$ and $\mathbf{w}$.

Solution for Problem 10.15:

(a) What's wrong with the following solution:

Bogus Solution: A 90° counterclockwise rotation of $\mathbf{i}$ is $\mathbf{j}$, and a 90° counterclockwise rotation of $\mathbf{j}$ is $-\mathbf{i}$. The product $\mathbf{Ni}$ equals the first column of $\mathbf{N}$, so the first column of $\mathbf{N}$ is $\mathbf{j}$. Similarly, the second column of $\mathbf{N}$ is $\mathbf{Nj}$, which equals $-\mathbf{i}$. Therefore, $\mathbf{N} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

This bogus solution does find the correct matrix. Moreover, the bogus solution shows that if the desired matrix exists, then $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ must be that matrix. However, the solution does not show that $\mathbf{Nv}$ is a 90° counterclockwise rotation of $\mathbf{v}$ *for all* $\mathbf{v}$.

Our first step in testing that $\mathbf{Nv}$ is a 90° counterclockwise rotation of $\mathbf{v}$ for all $\mathbf{v}$ is to figure out, in terms of x and y, what vector results when $\mathbf{v}$ is rotated 90° counterclockwise. We let P be (x, y) and P' be the image of P upon a 90° counterclockwise rotation. Rotation is often easier to analyze with polar coordinates. We let the polar coordinates of P be (r, θ), so the polar coordinates of P' are $(r, \theta + 90°)$. Converting these to rectangular coordinates, the rectangular coordinates of P are $x = r\cos\theta$ and $y = r\sin\theta$, and the rectangular coordinates of P' are

$$x' = r\cos(\theta + 90°) = r(\cos\theta\cos 90° - \sin\theta\sin 90°) = -r\sin\theta,$$
$$y' = r\sin(\theta + 90°) = r(\sin\theta\cos 90° + \sin 90°\cos\theta) = r\cos\theta.$$

Comparing our expressions for x' and y' to those for x and y, we find that $x' = -y$ and $y' = x$, so the rectangular coordinates of P' are $(-y, x)$.

Now we can check that the matrix we found in the Bogus Solution does indeed produce the desired result for all vectors $\mathbf{v}$:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}.$$

This matches our work above and thereby tells us that $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\mathbf{v}$ is indeed a 90° counterclockwise rotation of $\mathbf{v}$ for all $\mathbf{v}$.

(b) We showed earlier that the function represented by a matrix is a linear function. That is, we have $\mathbf{N}(c\mathbf{v}) = c\mathbf{Nv}$ and $\mathbf{N}(\mathbf{v} + \mathbf{w}) = \mathbf{Nv} + \mathbf{Nw}$ for any scalar c and any vectors $\mathbf{v}$ and $\mathbf{w}$. The fact that we can represent a 90° counterclockwise rotation of a vector as multiplication by a matrix $\mathbf{N}$ tells us that such a rotation is linear, too. □

Important: If a function that maps a vector to another vector can be represented by a matrix, then the function is linear.

We'll now generalize Problem 10.15.

Problem 10.16: Given an angle ϕ, find a matrix $\mathbf{T}$ such that for any vector $\mathbf{v}$, the vector $\mathbf{Tv}$ is a counterclockwise rotation of $\mathbf{v}$ by ϕ.

Solution for Problem 10.16: We use our solution for a 90° rotation as our guide. We let P be (x, y) and P' be the image of P upon a counterclockwise rotation by angle ϕ. Rotation suggests polar coordinates, so we let (r, θ) be the polar coordinates of P, which means that $x = r\cos\theta$, $y = r\sin\theta$, and the polar coordinates of P' are $(r, \theta + \phi)$. If (x', y') are the rectangular coordinates of P', then we must have

$$x' = r\cos(\theta + \phi) = r\cos\theta\cos\phi - r\sin\theta\sin\phi = x\cos\phi - y\sin\phi,$$
$$y' = r\sin(\theta + \phi) = r\sin\theta\cos\phi + r\cos\theta\sin\phi = y\cos\phi + x\sin\phi.$$

We could now find the columns of $\mathbf{T}$ by finding $\mathbf{Ti}$ and $\mathbf{Tj}$, and then confirm that the resulting matrix produces the desired expression for $\mathbf{Tv}$ for any $\mathbf{v}$. Or, we could simply consider the product $\mathbf{T}\begin{pmatrix} x \\ y \end{pmatrix}$ and compare it directly to the desired result. Taking the latter approach, we seek the matrix such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x\cos\phi - y\sin\phi \\ x\sin\phi + y\cos\phi \end{pmatrix}.$$

Therefore, we must have $ax + by = x\cos\phi - y\sin\phi$ and $cx + dy = x\sin\phi + y\cos\phi$, which means our desired matrix is

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}.$$

□

We can test that our matrix produces the correct result when $\phi = 0°$ and $\phi = 90°$. A rotation of 0° is the identity transformation—nothing is changed. Sure enough, setting $\phi = 0°$ gives us

$$\begin{pmatrix} \cos 0° & -\sin 0° \\ \sin 0° & \cos 0° \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

which is the identity matrix.

When $\phi = 90°$, we have

$$\begin{pmatrix} \cos 90° & -\sin 90° \\ \sin 90° & \cos 90° \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

which is the matrix we found in Problem 10.15 for a 90° counterclockwise rotation.

Important: If

$$\mathbf{T} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix},$$

then $\mathbf{Tv}$ is a counterclockwise rotation of $\mathbf{v}$ by an angle of ϕ.

Problem 10.17: A sequence (a_1, b_1), (a_2, b_2), (a_3, b_3), ... of points in the coordinate plane satisfies

$$(a_{n+1}, b_{n+1}) = (\sqrt{3}a_n - b_n, \sqrt{3}b_n + a_n)$$

for $n = 1, 2, 3, \ldots$. Suppose that $(a_{100}, b_{100}) = (2, 4)$. What is $a_1 + b_1$? *(Source: AMC 12)*

Solution for Problem 10.17: The ordered pair $(\sqrt{3}a_n - b_n,\ \sqrt{3}b_n + a_n)$ is a linear combination of the components of the ordered pair (a_n, b_n). This means that we can represent $\begin{pmatrix} \sqrt{3}a_n - b_n \\ \sqrt{3}b_n + a_n \end{pmatrix}$ as the product of a matrix and $\begin{pmatrix} a_n \\ b_n \end{pmatrix}$. Specifically, we have

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = \begin{pmatrix} \sqrt{3}a_n - b_n \\ \sqrt{3}b_n + a_n \end{pmatrix} = \begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix}.$$

The $\sqrt{3}$ entries make us think of rotation, but the entries clearly cannot all be cosines and sines because cosines and sines cannot be greater than 1. However, dividing all the entries in the matrix by 2 (factoring the scalar 2 out of the matrix) gives us some very familiar fractions:

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = 2\begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix}.$$

Now, we're in familiar territory, and we see that

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = 2\begin{pmatrix} \cos 30^\circ & -\sin 30^\circ \\ \sin 30^\circ & \cos 30^\circ \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix}.$$

The product $\begin{pmatrix} \cos 30^\circ & -\sin 30^\circ \\ \sin 30^\circ & \cos 30^\circ \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix}$ gives us a vector that is a 30° counterclockwise rotation of $\begin{pmatrix} a_n \\ b_n \end{pmatrix}$. Letting $\mathbf{T} = \begin{pmatrix} \cos 30^\circ & -\sin 30^\circ \\ \sin 30^\circ & \cos 30^\circ \end{pmatrix}$ and $\mathbf{v}_n = \begin{pmatrix} a_n \\ b_n \end{pmatrix}$, we have

$$\mathbf{v}_{100} = 2\mathbf{T}\mathbf{v}_{99} = 2\mathbf{T}(2\mathbf{T}\mathbf{v}_{98}) = \cdots = 2^{99}\underbrace{\mathbf{T}\cdots\mathbf{T}}_{99\ \mathbf{T}\text{'s}}\mathbf{v}_1.$$

Just as we use an exponent to denote a repeated product of any number, we can write the product of 99 $\mathbf{T}$'s as $\mathbf{T}^{99}$, so we have

$$\mathbf{v}_{100} = 2^{99}\mathbf{T}^{99}\mathbf{v}_1.$$

We know that $\mathbf{T}\mathbf{v}_1$ is a 30° counterclockwise rotation of $\mathbf{v}_1$, so $\mathbf{T}^{12}\mathbf{v}_1$ is a $12 \cdot 30^\circ = 360^\circ$ rotation of $\mathbf{v}_1$. Therefore, we have $\mathbf{T}^{12}\mathbf{v}_1 = \mathbf{v}_1$, which means that $\mathbf{T}^{12} = \mathbf{I}$. Since $99 = 3 + 12 \cdot 8$, we have

$$\mathbf{v}_{100} = 2^{99}\mathbf{T}^{99}\mathbf{v}_1 = 2^{99}\mathbf{T}^3\left(\mathbf{T}^{12}\right)^8\mathbf{v}_1 = 2^{99}\mathbf{T}^3\mathbf{v}_1.$$

Since $\mathbf{T}\mathbf{v}_1$ is a 30° counterclockwise rotation of $\mathbf{v}_1$, the product $\mathbf{T}^3\mathbf{v}_1$ gives us a vector that is a 90° counterclockwise rotation of $\mathbf{v}_1$. So, we have

$$\mathbf{v}_{100} = 2^{99}\begin{pmatrix} \cos 90^\circ & -\sin 90^\circ \\ \sin 90^\circ & \cos 90^\circ \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = 2^{99}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = 2^{99}\begin{pmatrix} -b_1 \\ a_1 \end{pmatrix},$$

which means $\begin{pmatrix} a_{100} \\ b_{100} \end{pmatrix} = \begin{pmatrix} -2^{99}b_1 \\ 2^{99}a_1 \end{pmatrix}$. We are given that $a_{100} = 2$ and $b_{100} = 4$, so

$$a_1 + b_1 = \frac{b_{100}}{2^{99}} - \frac{a_{100}}{2^{99}} = \frac{4-2}{2^{99}} = \frac{1}{2^{98}}.$$

□

Problem 10.18: Let $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$. Find the matrix $\mathbf{D}$ such that multiplying $\mathbf{v}$ by $\mathbf{D}$ corresponds to multiplying $x + yi$ by $a + bi$. (Your answer should be a matrix whose entries are in terms of a and b.)

Solution for Problem 10.18: We have

$$(a + bi)(x + yi) = (ax - by) + (bx + ay)i.$$

Therefore, we must have

$$\mathbf{D}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax - by \\ bx + ay \end{pmatrix}.$$

From this, we can easily read the components of $\mathbf{D}$:

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax - by \\ bx + ay \end{pmatrix}.$$

□

We've seen now that a great many different transformations of vectors can be represented by matrices. Now, we turn to our claim on page 338 that any linear function that maps vectors to vectors can be represented as multiplication by a matrix.

Problem 10.19: As a reminder, we say that a function f with domain $\mathbb{R}^2$ is a linear function if

(i) $f(a\mathbf{v}) = af(\mathbf{v})$ for all scalars a and vectors $\mathbf{v}$, and

(ii) $f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w})$ for all vectors $\mathbf{v}$ and $\mathbf{w}$.

Suppose $g(\mathbf{w})$ is a linear function that maps vectors to vectors.

(a) Explain why $g(x\mathbf{i} + y\mathbf{j}) = xg(\mathbf{i}) + yg(\mathbf{j})$.

(b) Why does part (a) tell us that there exists some matrix $\mathbf{A}$ such that $g(\mathbf{w}) = \mathbf{A}\mathbf{w}$ for all $\mathbf{w}$?

Solution for Problem 10.19:

(a) Applying item (ii) in our definition of a linear function, we have $g(x\mathbf{i} + y\mathbf{j}) = g(x\mathbf{i}) + g(y\mathbf{j})$. Then, we apply item (i) twice to find $g(x\mathbf{i}) + g(y\mathbf{j}) = xg(\mathbf{i}) + yg(\mathbf{j})$. Therefore, we have $g(x\mathbf{i} + y\mathbf{j}) = xg(\mathbf{i}) + yg(\mathbf{j})$.

(b) We can write any vector $\mathbf{v}$ as $v_1\mathbf{i} + v_2\mathbf{j}$ for some scalars v_1 and v_2. So, for any linear function g, applying part (a) gives us

$$g(\mathbf{v}) = g(v_1\mathbf{i} + v_2\mathbf{j}) = v_1g(\mathbf{i}) + v_2g(\mathbf{j}).$$

Since g maps vectors to vectors, there exist scalars a_{11}, a_{21}, a_{12}, and a_{22} such that $g(\mathbf{i}) = a_{11}\mathbf{i} + a_{21}\mathbf{j}$ and $g(\mathbf{j}) = a_{12}\mathbf{i} + a_{22}\mathbf{j}$. Therefore, we have

$$g(\mathbf{v}) = v_1(a_{11}\mathbf{i} + a_{21}\mathbf{j}) + v_2(a_{12}\mathbf{i} + a_{22}\mathbf{j}) = (v_1a_{11} + v_2a_{12})\mathbf{i} + (v_1a_{21} + v_2a_{22})\mathbf{j}.$$

If you don't see $\mathbf{A}\mathbf{v}$ on the right side in this notation, perhaps you will if we write the right side like this:

$$g(\mathbf{v}) = \begin{pmatrix} v_1a_{11} + v_2a_{12} \\ v_1a_{21} + v_2a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Letting $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, we have shown that there is a matrix $\mathbf{A}$ such that $g(\mathbf{v}) = \mathbf{A}\mathbf{v}$ for all $\mathbf{v}$. The columns of $\mathbf{A}$ are $g(\mathbf{i})$ and $g(\mathbf{j})$, so $\mathbf{A}$ is uniquely determined by g. Therefore, we have proved the following:

Important: Any linear function that maps vectors to vectors can be expressed as multiplication by a matrix, where the columns of the matrix are given by the vectors to which $\mathbf{i}$ and $\mathbf{j}$ are mapped by the function.

The final observation above gives us a very easy way to re-create the matrices we discovered in this chapter whenever we need them. For example, now that we know that rotation is a linear transformation, to find the matrix **N** that corresponds to a 90° counterclockwise rotation, we only need to find **Ni** and **Nj**. A 90° counterclockwise rotation of **i** is **j**, and a 90° counterclockwise rotation of **j** is **−i**. Therefore, the first column of **N** is **j** and the second is **−i**, which means $\mathbf{N} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Similarly, we can determine the effect of any linear transformation on any vector by examining the effect of the transformation on the easy-to-handle vectors **i** and **j**.

Exercises

10.3.1 Suppose that $\mathbf{A}\begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$. Must **A** be the identity matrix? If not, find a possible matrix **A** besides **I**.

10.3.2 Suppose a function g maps the vector **v** to the reflection of **v** over the x-axis. Is there a matrix **A** such that $g(\mathbf{v}) = \mathbf{Av}$ for all vectors **v**? If so, find it. If not, explain why there is not.

10.3.3 Suppose a function f maps the vector $\begin{pmatrix} x \\ y \end{pmatrix}$ to the vector $\begin{pmatrix} x^2 \\ y^2 \end{pmatrix}$. Is there a matrix **A** such that $f(\mathbf{v}) = \mathbf{Av}$ for all vectors **v**? If so, find it. If not, explain why there is not.

10.3.4 Show how to use the rotation matrix we discovered in Problem 10.16 to recall the angle sum identities for sine and cosine.

10.3.5 Suppose a point P is reflected over the x-axis and the image is rotated 90° counterclockwise about the origin. Is the result the same as if we first rotate P by 90° counterclockwise about the origin, and then reflect the result over the x-axis?

10.4 The Determinant

We introduced 2×2 matrices with a system of two linear equations with two variables. We start this section by considering when it is possible to find a unique solution to such a system.

Problem 10.20: Let a, b, c, d, u, and v be constants in the system of equations

$$ax + by = u,$$
$$cx + dy = v.$$

(a) Show that if $ad - bc \neq 0$, then the system has a unique solution (x, y).

(b) Show that if $ad - bc = 0$, then the system either has no solutions (x, y) or has infinitely many solutions.

Solution for Problem 10.20:

(a) We eliminate y by multiplying the first equation by d and the second by b to get

$$adx + bdy = du,$$
$$bcx + bdy = bv,$$

and then subtracting the second equation from the first to find

$$adx - bcx = du - bv.$$

Factoring the left gives $(ad - bc)x = du - bv$. Since $ad - bc \neq 0$, we can solve for x by dividing both sides to find

$$x = \frac{du - bv}{ad - bc}.$$

Similarly, we find y by subtracting c times $ax + by = u$ from a times $cx + dy = v$ to get

$$(ad - bc)y = av - cu.$$

Once again, we need to have $ad - bc \neq 0$ in order to divide both sides by $ad - bc$ to produce

$$y = \frac{av - cu}{ad - bc}.$$

Therefore, if $ad - bc \neq 0$, the system has the unique solution $(x, y) = \left(\frac{du-bv}{ad-bc}, \frac{av-cu}{ad-bc}\right)$.

(b) Before tackling the proof that there are either no solutions or infinitely many solutions, we'll get some geometric intuition for what's going on. The graphs of the two equations in the system are lines with direction vectors $\begin{pmatrix} -b \\ a \end{pmatrix}$ and $\begin{pmatrix} -d \\ c \end{pmatrix}$, respectively. These lines either don't intersect (and the system has no solutions) or are the same line (and the system has infinitely many solutions) if and only if $\begin{pmatrix} -b \\ a \end{pmatrix} = k\begin{pmatrix} -d \\ c \end{pmatrix}$ for some scalar k, or if $c = d = 0$. In either case, we have $ad - bc = 0$, since $b = kd$ and $a = kc$ gives $ad - bc = kcd - kcd = 0$, and $c = d = 0$ gives $ad - bc = 0$.

Now, let's take a more rigorous approach. In part (a), we showed that the system of equations implies the equations

$$(ad - bc)x = du - bv,$$
$$(ad - bc)y = av - cu.$$

If $ad - bc = 0$, then these equations become

$$0 = du - bv,$$
$$0 = av - cu.$$

So, if $du - bv \neq 0$ or $av - cu \neq 0$, then the corresponding equation above cannot be true, which means the system has no solutions. If $du - bv = 0$ and $av - cu = 0$, then we have $du = bv$ and $av = cu$. To see what this means for our original system of equations, $ax + by = u$, $cx + dy = v$, we multiply the first by v and the second by u, to get

$$avx + bvy = uv,$$
$$cux + duy = uv.$$

If $du = bv$ and $av = cu$, then these two equations are the same, so any of the infinitely many solutions (x, y) to the first equation are also solutions to the second. If u and v are nonzero, then dividing the first equation above by v and the second by u shows that these infinitely many values of (x, y) are solutions to the original system. With some careful casework, we can show that even if u or v is 0, then these infinitely many values of (x, y) are still solutions to the original system.

Therefore, if $ad - bc = 0$, the system

$$ax + by = u,$$
$$cx + dy = v,$$

either has no solutions or it has infinitely many solutions. □

We can write the original system of equations in Problem 10.20 as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}.$$

Our work in the solution shows that this system has a unique solution $\begin{pmatrix} x \\ y \end{pmatrix}$ if and only if $ad - bc \neq 0$. This is such an important property that for any matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we have a special name for the quantity $ad - bc$.

Definition: The **determinant** of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is the quantity $ad - bc$. We often denote this determinant as $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$, so

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

We also can denote the determinant of a matrix **A** as det(**A**).

In Problem 10.20, we proved that if $ad - bc \neq 0$, then the system of equations

$$\begin{aligned} ax + by &= u, \\ cx + dy &= v, \end{aligned}$$

has the unique solution $(x, y) = \left(\frac{du - bv}{ad - bc}, \frac{av - cu}{ad - bc} \right)$. We can write this solution in terms of determinants:

$$x = \frac{\begin{vmatrix} u & b \\ v & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}, \qquad y = \frac{\begin{vmatrix} a & u \\ c & v \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}.$$

If we express the system of equations as $\mathbf{M} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}$, where $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we see that each denominator in the solution to the system is det(**M**). The numerator of x is the determinant of the matrix formed by replacing the first column of **M** (which corresponds to the coefficients of x in the system of equations) with the column $\begin{pmatrix} u \\ v \end{pmatrix}$. Similarly, we find the numerator of y by replacing the second column of **M** with $\begin{pmatrix} u \\ v \end{pmatrix}$ and then taking the determinant of the resulting matrix. This method for solving a system of two linear equations with two variables is referred to as **Cramer's rule**.

For the rest of this section, we will explore properties of the determinant.

Problems

Problem 10.21: Solve each of the following systems of equations with Cramer's rule:

(a) $\begin{aligned} 4x - 14y &= 7, \\ -19x + 88y &= -1. \end{aligned}$

(b) $\begin{aligned} 9a &= 5b + 7, \\ -16b + 29a &= -3. \end{aligned}$

Problem 10.22:

(a) Evaluate $\begin{vmatrix} -2 & 3 \\ 4 & 7 \end{vmatrix}$, $\begin{vmatrix} -2+4 & 3+7 \\ 4 & 7 \end{vmatrix}$, and $\begin{vmatrix} -2+3\cdot 4 & 3+3\cdot 7 \\ 4 & 7 \end{vmatrix}$.

(b) Make a conjecture based on your results from part (a), and then prove your conjecture.

Problem 10.23:

(a) Suppose $\mathbf{B} = 2\mathbf{A}$. How are $\det(\mathbf{A})$ and $\det(\mathbf{B})$ related?

(b) Let $\mathbf{C}$ be the matrix formed by doubling the entries in the first row of $\mathbf{A}$. How are $\det(\mathbf{A})$ and $\det(\mathbf{C})$ related?

Problem 10.24: Evaluate $\begin{vmatrix} 157 & 199 \\ 314 & 399 \end{vmatrix}$ and $\begin{vmatrix} 33 & -77 \\ -1200 & 3600 \end{vmatrix}$.

Problem 10.25:

(a) Evaluate $\begin{vmatrix} 4 & 5 \\ -3 & 7 \end{vmatrix}$ and $\begin{vmatrix} -3 & 7 \\ 4 & 5 \end{vmatrix}$.

(b) Make a conjecture based on your results from part (a), and then prove your conjecture.

Problem 10.26:

(a) Find three matrices with nonzero components such that each has determinant 0.

(b) Notice anything interesting about your matrices in part (a)? Specifically, how are the rows related to each other in a matrix with determinant 0?

(c) How is your answer to part (b) related to what happens in a system of equations in which we cannot apply Cramer's rule because the denominators would come out to be 0?

Problem 10.27: Let $\mathbf{A} = \begin{pmatrix} 1 & 7 \\ -2 & -5 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 2 & -7 \\ 6 & 4 \end{pmatrix}$.

(a) Find $\det(\mathbf{A})$, $\det(\mathbf{B})$, $\det(\mathbf{AB})$, and $\det(\mathbf{A} + \mathbf{B})$.

(b) Make a conjecture based on your results from part (a), and then prove your conjecture.

We start with a little practice computing determinants and applying Cramer's rule.

Problem 10.21: Solve each of the following systems of equations with Cramer's rule:

(a) $\begin{aligned} 4x - 14y &= 7, \\ -19x + 88y &= -1. \end{aligned}$

(b) $\begin{aligned} 9a &= 5b + 7, \\ -16b + 29a &= -3. \end{aligned}$

Solution for Problem 10.21:

(a) Applying Cramer's rule gives

$$x = \frac{\begin{vmatrix} 7 & -14 \\ -1 & 88 \end{vmatrix}}{\begin{vmatrix} 4 & -14 \\ -19 & 88 \end{vmatrix}} = \frac{(7)(88) - (-1)(-14)}{(4)(88) - (-19)(-14)} = \frac{602}{86} = 7,$$

$$y = \frac{\begin{vmatrix} 4 & 7 \\ -19 & -1 \end{vmatrix}}{\begin{vmatrix} 4 & -14 \\ -19 & 88 \end{vmatrix}} = \frac{(4)(-1) - (-19)(7)}{86} = \frac{129}{86} = \frac{3}{2}.$$

(b) What's wrong with the following start:

Bogus Solution: Cramer's rule gives us

$$a = \frac{\begin{vmatrix} 7 & 5 \\ -3 & 29 \end{vmatrix}}{\begin{vmatrix} 9 & 5 \\ -16 & 29 \end{vmatrix}}, \qquad b = \frac{\begin{vmatrix} 9 & 7 \\ -16 & -3 \end{vmatrix}}{\begin{vmatrix} 9 & 5 \\ -16 & 29 \end{vmatrix}}.$$

There are all sorts of things wrong with this start, and they all stem from blindly sticking the coefficients and constants into Cramer's rule without first organizing the data in the problem in a way that makes Cramer's rule easy to apply. In our expression of Cramer's rule, we defined the coefficients such that all the variables are on the left sides of the equations and the constants are on the right sides. We also avoid mistakes by writing the variables in alphabetical order. Reorganizing the original system of equations in these ways gives us

$$\begin{aligned} 9a - 5b &= 7, \\ 29a - 16b &= -3. \end{aligned}$$

Now, we apply Cramer's rule to find

$$a = \frac{\begin{vmatrix} 7 & -5 \\ -3 & -16 \end{vmatrix}}{\begin{vmatrix} 9 & -5 \\ 29 & -16 \end{vmatrix}} = \frac{(7)(-16) - (-3)(-5)}{(9)(-16) - (29)(-5)} = \frac{-112 - 15}{-144 + 145} = -127,$$

$$b = \frac{\begin{vmatrix} 9 & 7 \\ 29 & -3 \end{vmatrix}}{\begin{vmatrix} 9 & -5 \\ 29 & -16 \end{vmatrix}} = \frac{(9)(-3) - (29)(7)}{(9)(-16) - (29)(-5)} = \frac{-27 - 203}{1} = -230.$$

□

Now that we have a little practice in computing determinants, let's explore some properties of determinants that can make evaluating them easier.

Problem 10.22:

(a) Evaluate $\begin{vmatrix} -2 & 3 \\ 4 & 7 \end{vmatrix}$, $\begin{vmatrix} -2+4 & 3+7 \\ 4 & 7 \end{vmatrix}$, and $\begin{vmatrix} -2+3\cdot 4 & 3+3\cdot 7 \\ 4 & 7 \end{vmatrix}$.

(b) Make a conjecture based on your results from part (a), and then prove your conjecture.

Solution for Problem 10.22:

(a) We have

$$\begin{aligned} \begin{vmatrix} -2 & 3 \\ 4 & 7 \end{vmatrix} &= (-2)(7) - (4)(3) = -26, \\ \begin{vmatrix} -2+4 & 3+7 \\ 4 & 7 \end{vmatrix} &= (-2+4)(7) - (4)(3+7) = -2(7) + 4(7) - 4(3) - 4(7) = -26, \\ \begin{vmatrix} -2+3\cdot 4 & 3+3\cdot 7 \\ 4 & 7 \end{vmatrix} &= (-2+3\cdot 4)(7) - (4)(3+3\cdot 7) = -2(7) + 3\cdot 4(7) - 4(3) - 4\cdot 3(7) = -26. \end{aligned}$$

The latter two matrices are formed by adding a constant times the second row to the first row. Our results suggest that this procedure will always leave the determinant unchanged. Our writing out the computations of the latter two determinants in great detail gives us a pretty good clue why this will always happen.

(b) Now that we know what to look for, it's easy to find. We wish to show that adding a multiple of one row to the other will not alter the determinant of the matrix. We have

$$\begin{vmatrix} a+kc & b+kd \\ c & d \end{vmatrix} = (a+kc)(d) - (c)(b+kd) = ad + kcd - bc - kcd = ad - bc = \begin{vmatrix} a & b \\ c & d \end{vmatrix}, \text{ and}$$

$$\begin{vmatrix} a & b \\ c+ka & d+kb \end{vmatrix} = (a)(d+kb) - (b)(c+ka) = ad + kab - bc - kab = ad - bc = \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Therefore, adding a multiple of one row of a 2×2 matrix to the other row of the matrix does not change the determinant of the matrix.

□

As an Exercise, you'll show that adding a multiple of one column of a 2×2 matrix to the other column of the matrix leaves the determinant of the matrix unchanged.

Important:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a+kc & b+kd \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c+ka & d+kb \end{vmatrix} = \begin{vmatrix} a+kb & b \\ c+kd & d \end{vmatrix} = \begin{vmatrix} a & b+ka \\ c & d+kc \end{vmatrix}.$$

Problem 10.23:

(a) Suppose $\mathbf{B} = k\mathbf{A}$. How are $\det(\mathbf{A})$ and $\det(\mathbf{B})$ related?

(b) Let $\mathbf{C}$ be the matrix formed by multiplying the entries in the first row of $\mathbf{A}$ by k. How are $\det(\mathbf{A})$ and $\det(\mathbf{C})$ related?

Solution for Problem 10.23: Let $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, so $\det(\mathbf{A}) = ad - bc$.

(a) We have

$$\det(\mathbf{B}) = \det(k\mathbf{A}) = \begin{vmatrix} ka & kb \\ kc & kd \end{vmatrix} = (ka)(kd) - (kb)(kc) = k^2(ad - bc) = k^2 \det(\mathbf{A}).$$

Important: For any 2×2 matrix $\mathbf{A}$ and any constant k, we have $\det(k\mathbf{A}) = k^2 \det(\mathbf{A})$.

(b) We have

$$\det(\mathbf{C}) = \begin{vmatrix} ka & kb \\ c & d \end{vmatrix} = (ka)(d) - (kb)(c) = k(ad - bc) = k \det(\mathbf{A}).$$

Similarly, multiplying the second row of $\mathbf{A}$ by any constant k gives a matrix with determinant $k \det(\mathbf{A})$, and multiplying either column of $\mathbf{A}$ by any constant k gives a matrix with determinant $k \det(\mathbf{A})$.

Important: Multiplying either column or either row of a matrix $\mathbf{A}$ by any constant k gives a matrix with determinant $k \det(\mathbf{A})$.

□

We can sometimes use the rules that we proved in the previous two problems to simplify the computation of the determinant of a matrix.

Problem 10.24: Evaluate $\begin{vmatrix} 157 & 199 \\ 314 & 399 \end{vmatrix}$ and $\begin{vmatrix} 33 & -77 \\ -1200 & 3600 \end{vmatrix}$.

Solution for Problem 10.24: For the first matrix, we notice that 314 is double 157, so we add -2 times the first row to the second row, and we have

$$\begin{vmatrix} 157 & 199 \\ 314 & 399 \end{vmatrix} = \begin{vmatrix} 157 & 199 \\ 314 - 2(157) & 399 - 2(199) \end{vmatrix} = \begin{vmatrix} 157 & 199 \\ 0 & 1 \end{vmatrix} = (157)(1) - (0)(199) = 157.$$

Turning to the second matrix, we see a common factor of 11 in the first row and 1200 in the second. Repeatedly applying the principle we proved in part (b) of Problem 10.23, we have

$$\begin{vmatrix} 33 & -77 \\ -1200 & 3600 \end{vmatrix} = 11 \begin{vmatrix} 3 & -7 \\ -1200 & 3600 \end{vmatrix} = (11)(1200) \begin{vmatrix} 3 & -7 \\ -1 & 3 \end{vmatrix} = (11)(1200)((3)(3) - (-7)(-1)) = 26400.$$

□

Problem 10.25:

(a) Evaluate $\begin{vmatrix} 4 & 5 \\ -3 & 7 \end{vmatrix}$ and $\begin{vmatrix} -3 & 7 \\ 4 & 5 \end{vmatrix}$.

(b) Make a conjecture based on your results from part (a), and then prove your conjecture.

Solution for Problem 10.25:

(a) We have $\begin{vmatrix} 4 & 5 \\ -3 & 7 \end{vmatrix} = (4)(7) - (-3)(5) = 28 + 15 = 43$ and $\begin{vmatrix} -3 & 7 \\ 4 & 5 \end{vmatrix} = (-3)(5) - (4)(7) = -15 - 28 = -43.$

(b) It looks like swapping the rows of a 2×2 matrix $\mathbf{A}$ produces a matrix with determinant $(-1)\det(\mathbf{A})$. Proving this is true is straightforward:

$$\begin{vmatrix} c & d \\ a & b \end{vmatrix} = cb - ad = -(ad - bc) = -\begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

□

Important: For any 2×2 matrix $\mathbf{A}$, swapping the rows or columns of $\mathbf{A}$ produces a matrix with determinant $(-1)\det(\mathbf{A})$.

Problem 10.26:

(a) Find three matrices with nonzero components such that each has determinant 0.

(b) Notice anything interesting about your matrices in part (a)? Specifically, how are the rows related to each other in a matrix with determinant 0?

(c) How is your answer to part (b) related to what happens in a system of equations in which we cannot apply Cramer's rule because the denominators would come out to be 0?

Solution for Problem 10.26:

(a) With a little trial and error, we can find many matrices with determinant 0. Here are some:

$$\begin{vmatrix} 1 & 1 \\ -1 & -1 \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ 2 & 3 \end{vmatrix} = \begin{vmatrix} 3 & -5 \\ 6 & -10 \end{vmatrix} = \begin{vmatrix} -2 & -8 \\ 1 & 4 \end{vmatrix} = \begin{vmatrix} -5 & 6 \\ 25 & -30 \end{vmatrix} = 0.$$

(b) In each of the example matrices in part (a), the first row is a multiple of the second. To see why this is happening, we note that $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$. This equals 0 if and only if $ad = bc$, which we can write as $\frac{a}{c} = \frac{b}{d}$

when $cd \neq 0$. When $\frac{a}{c} = \frac{b}{d}$, the first row of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a multiple of the second. If $c = 0$, then we can only have $ad - bc = 0$ if $a = 0$ or $d = 0$. In either case, the second row is a multiple of the first. (Similarly, if $d = 0$, we have $bc = 0$, from which we find that the second row is a multiple of the first.)

Therefore, the determinant of a 2×2 matrix is 0 if and only if one row is a multiple of the other.

(c) We can only apply Cramer's rule to the system

$$ax + by = u,$$
$$cx + dy = v,$$

if $ad - bc \neq 0$. What goes wrong when $ad - bc = 0$? As we saw in part (b), we have $ad - bc = 0$ if and only if one row of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a multiple of the other. This corresponds to the left side of one of the equations in our system being a constant multiple of the other left side. Geometrically speaking, the graphs of the two equations are either the same line or parallel lines. So, we cannot have a unique solution to the system of equations.

□

Just as we can describe a pair of vectors as being linearly dependent if one is a constant times the other, we can describe two rows of a matrix as being linearly dependent if either is a constant times the other.

Important: The determinant of a 2×2 matrix is 0 if and only if the rows of the matrix are linearly dependent.

We can equivalently state this fact as, "The determinant of a 2×2 matrix is nonzero if and only if the rows of the matrix are linearly independent."

Matrices with determinant 0 are so important that we have a special name for them. A matrix with determinant 0 is called a **singular matrix**, and a matrix with nonzero determinant is called a **nonsingular matrix.**

Problem 10.27: Let $\mathbf{A} = \begin{pmatrix} 1 & 7 \\ -2 & -5 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 2 & -7 \\ 6 & 4 \end{pmatrix}$.

(a) Find $\det(\mathbf{A})$, $\det(\mathbf{B})$, $\det(\mathbf{AB})$, and $\det(\mathbf{A} + \mathbf{B})$.

(b) Make a conjecture based on your results from part (a), and then prove your conjecture.

Solution for Problem 10.27:

(a) We have $\det(\mathbf{A}) = (1)(-5) - (7)(-2) = -5 + 14 = 9$, $\det(\mathbf{B}) = (2)(4) - (-7)(6) = 8 + 42 = 50$, $\det(\mathbf{A} + \mathbf{B}) = \begin{vmatrix} 3 & 0 \\ 4 & -1 \end{vmatrix} = -3$, and

$$\det(\mathbf{AB}) = \begin{vmatrix} 44 & 21 \\ -34 & -6 \end{vmatrix} = 2\begin{vmatrix} 44 & 21 \\ -17 & -3 \end{vmatrix} = 2 \cdot 3\begin{vmatrix} 44 & 7 \\ -17 & -1 \end{vmatrix} = 6((44)(-1) - (7)(-17)) = 6(75) = 450.$$

(b) In part (a), we don't see any obvious relationship between $\det(\mathbf{A} + \mathbf{B})$ and $\det(\mathbf{A}) + \det(\mathbf{B})$, but we do have $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$. Let's see if this is always the case. We have

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} &= (a_{11}a_{22} - a_{21}a_{12})(b_{11}b_{22} - b_{21}b_{12}) \\ &= a_{11}a_{22}b_{11}b_{22} - a_{11}a_{22}b_{21}b_{12} - a_{21}a_{12}b_{11}b_{22} + a_{21}a_{12}b_{21}b_{12} \\ &= a_{11}a_{22}b_{11}b_{22} + a_{12}a_{21}b_{12}b_{21} - a_{11}a_{22}b_{12}b_{21} - a_{12}a_{21}b_{11}b_{22}, \end{aligned}$$

where we have reorganized the terms in the last row to make it easier to compare to det(**AB**). Speaking of which, we have

$$\begin{aligned}\det\left(\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{pmatrix}\right) &= \begin{vmatrix}a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22}\\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22}\end{vmatrix}\\ &= (a_{11}b_{11}+a_{12}b_{21})(a_{21}b_{12}+a_{22}b_{22})-(a_{11}b_{12}+a_{12}b_{22})(a_{21}b_{11}+a_{22}b_{21})\\ &= a_{11}b_{11}a_{21}b_{12}+a_{11}b_{11}a_{22}b_{22}+a_{12}b_{21}a_{21}b_{12}+a_{12}b_{21}a_{22}b_{22}\\ &\quad -a_{11}b_{12}a_{21}b_{11}-a_{11}b_{12}a_{22}b_{21}-a_{12}b_{22}a_{21}b_{11}-a_{12}b_{22}a_{22}b_{21}.\end{aligned}$$

We can clean this up a bit by organizing the terms in each product. When we do so, we see that we have some convenient cancellation:

$$\begin{aligned}\det\left(\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{pmatrix}\right) &= a_{11}a_{21}b_{11}b_{12}+a_{11}a_{22}b_{11}b_{22}+a_{12}a_{21}b_{12}b_{21}+a_{12}a_{22}b_{21}b_{22}\\ &\quad -a_{11}a_{21}b_{11}b_{12}-a_{11}a_{22}b_{12}b_{21}-a_{12}a_{21}b_{11}b_{22}-a_{12}a_{22}b_{21}b_{22}\\ &= a_{11}a_{22}b_{11}b_{22}+a_{12}a_{21}b_{12}b_{21}-a_{11}a_{22}b_{12}b_{21}-a_{12}a_{21}b_{11}b_{22},\end{aligned}$$

which does indeed match our expression above for det(**A**) det(**B**).

□

Important: We have det(**AB**) = det(**A**) det(**B**) for any 2×2 matrices **A** and **B**.

WARNING!! The values of det(**A**) + det(**B**) and det(**A** + **B**) are *not* necessarily equal.

Exercises

10.4.1 Use Cramer's rule to solve the system $3x + 4y = -3x + 7$, $5y = -2x + 9$.

10.4.2 Evaluate the following determinants:

(a) $\begin{vmatrix}-3 & -5\\ 6 & 18\end{vmatrix}$

(b) $\begin{vmatrix}2400 & 3200\\ 25/3 & 31/3\end{vmatrix}$

10.4.3 Is it true that det(**AB**) = det(**BA**)?

10.4.4 The first column of **B** is the same as the first column of **A**. The second column of **B** is the sum of the second column of **A** and k times the first column of **A**. Show that det(**A**) = det(**B**).

10.4.5 We showed in the text that det(**A**) + det(**B**) does not necessarily equal det(**A** + **B**).

(a) What if det(**A**) = det(**B**) = 0? Then, must we have det(**A**) + det(**B**) = det(**A** + **B**)?

(b) What if we have det(**A**) = 0 but det(**B**) ≠ 0. Then, must we have det(**A**) + det(**B**) = det(**A** + **B**)?

10.5 Geometric Interpretation of the Determinant

Problems

Problem 10.28: In this problem, we investigate the geometric significance of the determinant of a 2×2 matrix. Let $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, let O be the origin, and let V and W be the points such that $\vec{V} = \mathbf{Mi}$ and $\vec{W} = \mathbf{Mj}$.

(a) Show that the area of the parallelogram with $\overline{OV}$ and $\overline{OW}$ as sides equals $\|\vec{V}\|\|\vec{W}\| \sin\theta$, where θ is the angle between $\vec{V}$ and $\vec{W}$.

(b) Express the area in terms of $\vec{V} \cdot \vec{W}$, $\|\vec{V}\|$, and $\|\vec{W}\|$.

(c) Find the area in terms of a, b, c, and d.

Problem 10.29:

(a) What is the geometric significance of a 0 determinant?

(b) Find several matrices with a negative determinant, and draw the corresponding parallelograms. Make a conjecture about the geometric significance of a negative determinant.

(c) Prove your conjecture in part (b). **Hints:** 160, 112

Problem 10.30: Use the geometric interpretation of the determinant to explain why $\det(k\mathbf{A}) = k^2 \det(\mathbf{A})$ for any 2×2 matrix $\mathbf{A}$.

Problem 10.31: Use the geometric interpretation of the determinant to explain why adding a multiple of either column of a 2×2 matrix to the other column leaves the determinant of the matrix unchanged.

Problem 10.32: In this problem, we find a method for finding the area of a triangle given the coordinates of its vertices.

(a) Find the area of a triangle with vertices $(0,0)$, $(5,4)$, and $(1,-3)$.

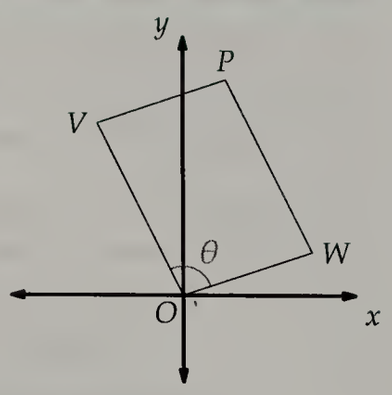
(b) Find the area of a triangle with vertices $(4,5)$, $(7,10)$, and $(-3,7)$.

(c) Find a formula for the area of a triangle with vertices (x_1, y_1), (x_2, y_2), and (x_3, y_3), where these vertices are in counterclockwise order.

Problem 10.28: Let O be the origin, V be (a,c), and W be (b,d). Show that the area of the parallelogram with $\overline{OV}$ and $\overline{OW}$ as sides is the absolute value of $\det(\mathbf{M})$, where $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Solution for Problem 10.28: First, we find the fourth vertex of the parallelogram by noting that if P is the point such that $\vec{V} + \vec{W} = \vec{P}$, then $OVPW$ is a parallelogram. Since $\triangle OVW \cong \triangle PWV$, we have $[OVPW] = 2[OVW] = 2\left(\frac{1}{2}(OV)(OW)\sin\theta\right)$, where $\theta = \angle VOW$. Therefore, we have $[OVPW] = \|\vec{V}\|\|\vec{W}\| \sin\theta$, where θ is the angle between $\vec{V}$ and $\vec{W}$. The dot product gives us another relationship involving $\vec{V}$, $\vec{W}$, and θ:

$$\vec{V} \cdot \vec{W} = \|\vec{V}\|\|\vec{W}\| \cos\theta.$$



We can now eliminate $\sin\theta$ from our expression for $[OVPW]$. Since $0° \le \theta \le 180°$, we have $\sin\theta \ge 0$, so we take the positive root from $\sin\theta = \sqrt{1-\cos^2\theta}$, and we have

$$\begin{aligned}
[OVPW] &= \|\vec{V}\|\|\vec{W}\|\sin\theta \\
&= \|\vec{V}\|\|\vec{W}\|\sqrt{1-\cos^2\theta} \\
&= \|\vec{V}\|\|\vec{W}\|\sqrt{1-\frac{(\vec{V}\cdot\vec{W})^2}{\|\vec{V}\|^2\|\vec{W}\|^2}} \\
&= \sqrt{\|\vec{V}\|^2\|\vec{W}\|^2-(\vec{V}\cdot\vec{W})^2} \\
&= \sqrt{(a^2+c^2)(b^2+d^2)-(ab+cd)^2} \\
&= \sqrt{a^2b^2+a^2d^2+b^2c^2+c^2d^2-a^2b^2-2abcd-c^2d^2} \\
&= \sqrt{a^2d^2-2abcd+b^2c^2} = \sqrt{(ad-bc)^2} = |ad-bc|,
\end{aligned}$$

as desired. □

Important: The absolute value of the determinant $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ equals the area of the parallelogram that has as two sides the vectors $\begin{pmatrix} a \\ c \end{pmatrix}$ and $\begin{pmatrix} b \\ d \end{pmatrix}$ drawn from the origin.

Repeatedly referring to the "absolute value" of det(**M**) is pretty unwieldy. Let's take a closer look at what a negative determinant corresponds to geometrically.

Problem 10.29:

(a) What is the geometric significance of a 0 determinant?

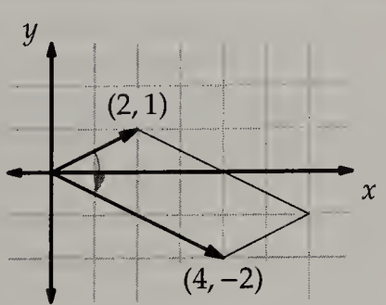
(b) Find several matrices with a negative determinant, and draw the corresponding parallelograms. Make a conjecture about the geometric significance of a negative determinant.

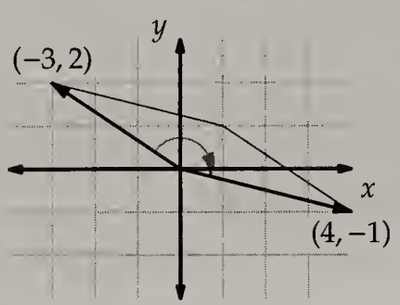
(c) Prove your conjecture in part (b).

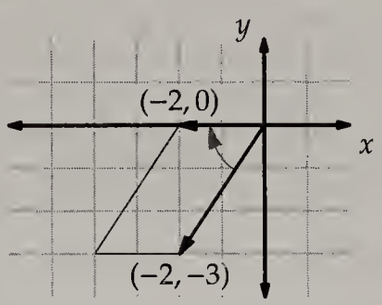
Solution for Problem 10.29:

(a) If **v** and **w** are nonzero, then the parallelogram with sides **v** and **w** drawn from the origin has area 0 if and only if **v** and **w** are in the same or opposite directions. This occurs if and only if **v** is a multiple of **w**. This corresponds to the fact that the determinant $\begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix}$ is zero if and only if the columns are linearly dependent.

(b) To learn the significance of a negative determinant, we examine a few cases. Below are the parallelograms whose areas correspond to the determinants $\begin{vmatrix} 2 & 4 \\ 1 & -2 \end{vmatrix}$, $\begin{vmatrix} -3 & 4 \\ 2 & -1 \end{vmatrix}$, and $\begin{vmatrix} -2 & -2 \\ -3 & 0 \end{vmatrix}$.

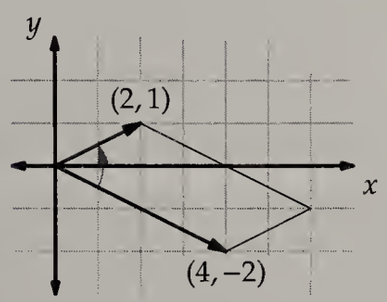


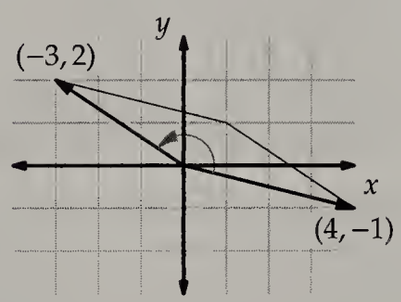


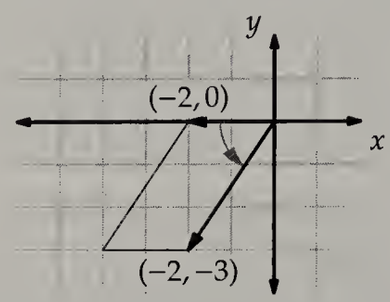


In each case, the determinant is negative. In each case, the smallest rotation from the vector given by the first column to the vector given by the second column is a *clockwise* rotation, as indicated by the arrows in the diagrams.

Swapping the columns of a 2×2 matrix produces a matrix whose determinant is the opposite of the determinant of the original matrix. Doing so for our three example matrices gives us $\begin{vmatrix} 4 & 2 \\ -2 & 1 \end{vmatrix}$, $\begin{vmatrix} 4 & -3 \\ -1 & 2 \end{vmatrix}$, and $\begin{vmatrix} -2 & -2 \\ 0 & -3 \end{vmatrix}$, and the corresponding parallelograms are below.







These are the same parallelograms as before, since the columns of the matrices are the same as before, only in opposite order. However, now the smallest rotation from the direction of the first column to the direction of the second is *counterclockwise*. We conjecture that the determinant of a 2×2 matrix is positive if and only if the smallest rotation from the vector of the first column to the vector of the second is counterclockwise.

(c) Let the columns of $\mathbf{A}$ be nonzero vectors $\mathbf{v}$ and $\mathbf{w}$, so $\mathbf{A} = \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix}$. (If $\mathbf{v}$ or $\mathbf{w}$ is $\mathbf{0}$, then $\det(\mathbf{A}) = 0$.) We need to relate $\det(\mathbf{A})$ to the angle between $\mathbf{v}$ and $\mathbf{w}$, so we let θ be the smallest rotation that must be applied to $\mathbf{v}$ to make it have the same direction as $\mathbf{w}$. (As usual, if $\theta > 0$, this rotation is counterclockwise, and if $\theta < 0$, then this rotation is clockwise.) Therefore, the vector $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \mathbf{v}$ has the same direction as $\mathbf{w}$, so there is some positive constant k such that

$$\mathbf{w} = k \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \mathbf{v} = k \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} kv_1\cos\theta - kv_2\sin\theta \\ kv_1\sin\theta + kv_2\cos\theta \end{pmatrix}.$$

(Specifically, $k = \|\mathbf{w}\| \,/\, \|\mathbf{v}\|$.)

Therefore, we have

$$\begin{aligned} \det(\mathbf{A}) &= \begin{vmatrix} v_1 & kv_1\cos\theta - kv_2\sin\theta \\ v_2 & kv_1\sin\theta + kv_2\cos\theta \end{vmatrix} \\ &= v_1(kv_1\sin\theta + kv_2\cos\theta) - v_2(kv_1\cos\theta - kv_2\sin\theta) \\ &= k(v_1^2 + v_2^2)\sin\theta. \end{aligned}$$

Aha! Now we have the whole story. Since k, v_1^2, and v_2^2 are all nonnegative, the sign of $\det(\mathbf{A})$ must match the sign of $\sin\theta$. We have $\sin\theta > 0$ for $0 < \theta < \pi$, and $\sin\theta < 0$ for $-\pi < \theta < 0$. Since positive values of θ correspond to counterclockwise rotations and negative values of θ correspond to clockwise rotations, we see that the determinant of a matrix is negative if the smallest rotation from the direction of the first column to the direction of the second column is a *clockwise* rotation, and positive if the smallest such rotation is a *counterclockwise* rotation. If the two columns have the same or opposite directions, then the columns are linearly dependent and the determinant is 0.

□

We can get rid of the "absolute value" in the statement, "The absolute value of the determinant $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ equals the area of the parallelogram that has as two sides the vectors $\begin{pmatrix} a \\ c \end{pmatrix}$ and $\begin{pmatrix} b \\ d \end{pmatrix}$ drawn from the origin" by using the concept of **signed area**. We define the signed area of a polygon by first assigning an orientation, clockwise or counterclockwise, to the polygon. If the orientation is counterclockwise, as shown in the diagram at left below, then we say the signed area is positive. If the orientation is clockwise, then we say the signed area is negative.

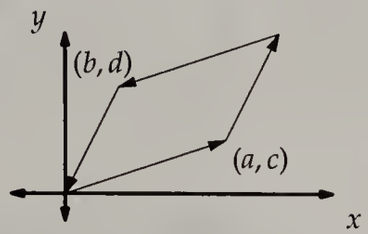


Figure 10.1: A Parallelogram with Positive Area

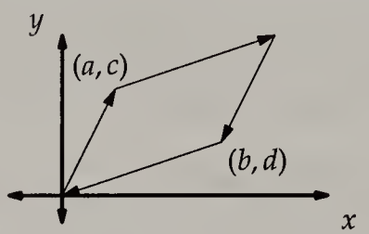


Figure 10.2: A Parallelogram with Negative Area

With this definition in hand, we can refine the relationship between the determinant of a matrix and area.

Important: The value of the determinant $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ equals the signed area of the parallelogram that has as two sides the vectors $\begin{pmatrix} a \\ c \end{pmatrix}$ and $\begin{pmatrix} b \\ d \end{pmatrix}$ drawn from the origin, where the orientation of the parallelogram is determined by the path around the parallelogram with first step from the origin to (a, c).

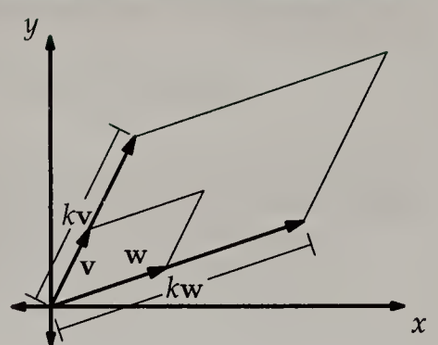
We learned in Section 10.3 that multiplying a vector by a matrix corresponds to a transformation of the vector. Our work so far in this section tells us about the effect of such a transformation on area.

Specifically, we showed that if $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and we let $\mathbf{v} = \mathbf{Mi} = \begin{pmatrix} a \\ c \end{pmatrix}$ and $\mathbf{w} = \mathbf{Mj} = \begin{pmatrix} b \\ d \end{pmatrix}$, then the signed area of the parallelogram with sides $\mathbf{v}$ and $\mathbf{w}$ is $\det(\mathbf{M})$. Since the signed area of the parallelogram with sides $\mathbf{i}$ and $\mathbf{j}$ is simply 1, we see that multiplication by $\mathbf{M}$ scales the area of this parallelogram by $\det(\mathbf{M})$.

We can extend this result to any two linearly independent vectors, not just $\mathbf{i}$ and $\mathbf{j}$. If $\mathbf{r}$ and $\mathbf{s}$ are linearly independent, then the signed area of the parallelogram with sides $\mathbf{Mr}$ and $\mathbf{Ms}$ is $\det(\mathbf{M})$ times the signed area of the parallelogram with sides $\mathbf{r}$ and $\mathbf{s}$. You'll be asked to prove this claim as a Challenge Problem.

Problem 10.30: Use the geometric interpretation of the determinant to explain why $\det(k\mathbf{A}) = k^2 \det(\mathbf{A})$ for any 2×2 matrix $\mathbf{A}$.

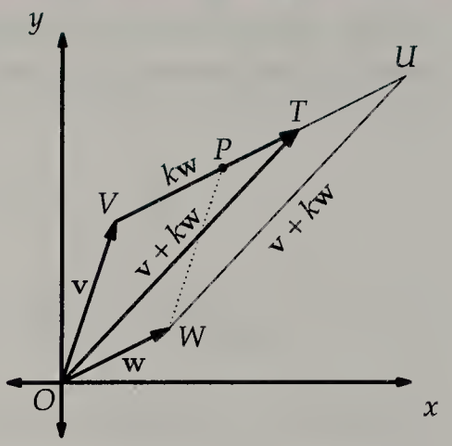
Solution for Problem 10.30: Let $\mathcal{P}$ be the parallelogram with sides $\mathbf{v}$ and $\mathbf{w}$ drawn from the origin, and let $\mathcal{Q}$ be the parallelogram with sides $k\mathbf{v}$ and $k\mathbf{w}$ drawn from the origin. $\mathcal{Q}$ is formed by scaling the sides of $\mathcal{P}$ by a factor of k. Therefore, $\mathcal{Q}$ and $\mathcal{P}$ are similar, and the ratio of the lengths of corresponding sides is k, which means the ratio of their areas is k^2.



In our discussion above, we've glossed over the issue of negative determinants and what happens when $k < 0$. Our discussion in Problem 10.29 takes care of both these concerns. The smallest rotation from the direction of $\mathbf{v}$ to the direction of $\mathbf{w}$ is the same as the smallest rotation from the direction of $k\mathbf{v}$ to the direction of $k\mathbf{w}$ (no matter what the sign of k is). Therefore, the determinants of $\mathbf{A}$ and $k\mathbf{A}$ must have the same sign. □

Problem 10.31: Use the geometric interpretation of the determinant to explain why adding a multiple of either column of a 2×2 matrix to the other column leaves the determinant of the matrix unchanged.

Solution for Problem 10.31: To add $k\mathbf{w}$ to $\mathbf{v}$ geometrically, we first draw $\mathbf{v}$ from the origin to V, and then draw $k\mathbf{w}$ starting from V. Because $k\mathbf{w}$ is parallel to $\mathbf{w}$, the head of $\mathbf{v} + k\mathbf{w}$, point T in the diagram, is on the line through V parallel to $\mathbf{w}$. Therefore, the distances from V and T to $\overleftrightarrow{OW}$ are equal, which means that the areas of OWV and OWT are the same. This means that the area of the parallelogram with sides $\mathbf{v}$ and $\mathbf{w}$ (parallelogram $OVPW$) equals the area of the parallelogram with sides $\mathbf{v} + k\mathbf{w}$ and $\mathbf{w}$ (parallelogram $OTUW$).



Letting $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$, this gives us

$$\begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix} = \begin{vmatrix} v_1 + kw_1 & w_1 \\ v_2 + kw_2 & w_2 \end{vmatrix}.$$

We haven't yet explained why the signs of these determinants are the same. When determining the orientation of the parallelogram corresponding to each of the determinants above, the last step in the path around the parallelogram in each is from W to O. So, the two parallelograms have the same orientation if these last steps complete circuits about the parallelograms with the same orientation. This is the case if and only if V and T are not on opposite sides of $\overleftrightarrow{OW}$. Because $\overline{VT} \parallel \overline{OW}$, we know that V and T (the heads of $\mathbf{v}$ and $\mathbf{v} + k\mathbf{w}$) cannot be on opposite sides of $\overleftrightarrow{OW}$. Therefore, the orientations of the parallelograms corresponding to the determinants are the same, so the signs of the determinants are the same. (In the case shown, both are clockwise, and hence have negative area.)

Similarly, we can add any multiple of the first column to the second without changing the determinant. □

Problem 10.32: In this problem, we find a method for finding the area of a triangle given the coordinates of its vertices.

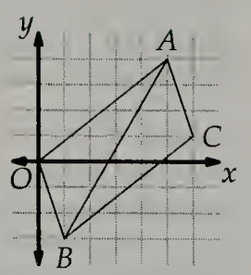
(a) Find the area of a triangle with vertices $O = (0,0)$, $A = (5,4)$, and $B = (1,-3)$.

(b) Find the area of a triangle with vertices $P = (4,5)$, $Q = (7,10)$, and $R = (-3,7)$.

(c) Find a formula for the area of a triangle with vertices (x_1, y_1), (x_2, y_2), and (x_3, y_3), where these vertices are in counterclockwise order.

Solution for Problem 10.32:

(a) We know how to handle a parallelogram, so we "complete" the parallelogram by finding the point C such that $\vec{C} = \vec{A} + \vec{B}$. As shown, this point is $(5+1, 4+(-3)) = (6,1)$. Since $\triangle OAB \cong \triangle CBA$, we have



$$[OAB] = \frac{1}{2}[OACB] = \frac{1}{2}\left|\det \begin{pmatrix} 5 & 1 \\ 4 & -3 \end{pmatrix}\right| = \frac{1}{2}|(5)(-3) - (4)(1)| = \frac{19}{2}.$$

(b) From part (a) we know how to find the area of a triangle with one vertex at the origin, so we translate the triangle so that P is at the origin. The translation that maps $(4,5)$ to $(0,0)$ also maps $(7,10)$ to $(7-4, 10-5) = (3,5)$ and maps $(-3,7)$ to $(-3-4, 7-5) = (-7,2)$. Therefore, we now seek the area of the triangle with vertices $(0,0)$, $(3,5)$ and $(-7,2)$. By the same argument as in part (a), the area of this triangle is

$$\frac{1}{2}\left|\det \begin{pmatrix} 3 & -7 \\ 5 & 2 \end{pmatrix}\right| = \frac{1}{2}|(3)(2) - (-7)(5)| = \frac{41}{2}.$$

(c) Part (b) gives us a guide; we start by translating one vertex to the origin. The same translation that maps (x_1, y_1) to $(0,0)$ also maps (x_2, y_2) to $(x_2 - x_1, y_2 - y_1)$ and maps (x_3, y_3) to $(x_3 - x_1, y_3 - y_1)$. Since the vertices (x_1, y_1), (x_2, y_2), and (x_3, y_3) are in counterclockwise order, so are the vertices $(0,0)$, $(x_2 - x_1, y_2 - y_1)$, and $(x_3 - x_1, y_3 - y_1)$. The area of this triangle is half the area of the parallelogram with $\begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}$ and $\begin{pmatrix} x_3 - x_1 \\ y_3 - y_1 \end{pmatrix}$ as sides, so the desired area is

$$\frac{1}{2} \det \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix}.$$

Note that if the vertices were in clockwise order rather than counterclockwise, then the expression above is negative, and gives the signed area of the triangle.

□

Exercises

10.5.1 Find the determinant of the rotation matrix $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Explain the result geometrically.

10.5.2 Find the area of the quadrilateral with vertices $(-2,3)$, $(1,-5)$, $(7,-4)$, and $(8,1)$.

10.5.3 Explain geometrically why multiplying one column of $\mathbf{A}$ by a constant k produces a matrix with determinant $k \det(\mathbf{A})$.

10.5.4 A **lattice point** is a point in the Cartesian plane with integer coordinates. Suppose $\mathcal{T}$ is a triangle in the Cartesian plane such that all three vertices of $\mathcal{T}$ are lattice points. Show that doubling the area of $\mathcal{T}$ must produce an integer.

10.5.5★ In Exercise 10.1.3, we discovered a matrix $\mathbf{F}$ such that multiplying a matrix on the right by $\mathbf{F}$ swaps the columns of the matrix. That is, the first column of $\mathbf{AF}$ is the second column of $\mathbf{A}$, and the second column of $\mathbf{AF}$ is the first column of $\mathbf{A}$. Explain geometrically what the determinant of $\mathbf{F}$ must be.

10.6 Inverting a Matrix

We have seen that we can solve the system of equations described by

$$\begin{pmatrix} 3 & -4 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ -10 \end{pmatrix}$$

with Cramer's rule. Rather than grinding through Cramer's rule, it sure would be nice to be able to divide by the matrix on the left. Unfortunately, division by a matrix is not defined. However, we can take another look solving equations with division by real numbers to get an idea how we might accomplish our goal. Consider the equation

$$3x = 9.$$

To solve this equation we can "divide both sides by 3." What we're really doing when we "divide both sides by 3" is "multiplying both sides by $\frac{1}{3}$." More precisely, we are multiplying by the multiplicative inverse of 3, which eliminates the 3 on the left side. Multiplying both sides of $3x = 9$ by $\frac{1}{3}$ gives

$$\frac{1}{3} \cdot 3x = \frac{1}{3} \cdot 9.$$

The $\frac{1}{3}$ and 3 cancel, and we're left with $x = \frac{1}{3} \cdot 9 = 3$.

Let's see if we can accomplish something similar with matrices.

Problems

Problem 10.33: Suppose that $\mathbf{ABv} = \mathbf{v}$ for all vectors $\mathbf{v}$. What matrix must $\mathbf{AB}$ equal?

Problem 10.34: If $\mathbf{AB}$ equals the answer you found in Problem 10.33, then we say that $\mathbf{A}$ is the **inverse** of $\mathbf{B}$ and that $\mathbf{B}$ is **invertible**. We often denote the inverse of a matrix $\mathbf{B}$ as $\mathbf{B}^{-1}$. Let $\mathbf{M} = \begin{pmatrix} 2 & 3 \\ 6 & 7 \end{pmatrix}$. Find $\mathbf{M}^{-1}$.

Problem 10.35: Does the matrix $\mathbf{M} = \begin{pmatrix} 3 & 5 \\ -6 & -10 \end{pmatrix}$ have an inverse? If it does, then find it. If it doesn't, explain why not.

Problem 10.36: If $\mathbf{A}$ is invertible, then how are $\det(\mathbf{A})$ and $\det(\mathbf{A}^{-1})$ related?

Problem 10.37: Let $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. In this problem, we find the conditions under which $\mathbf{A}$ is invertible, and we find the inverse in terms of a, b, c, and d.

(a) Suppose $\mathbf{A}$ is invertible, and let $\mathbf{A}^{-1} = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$. What matrix must $\mathbf{A}^{-1}\mathbf{A}$ equal?

(b) Use your answer to (a) to find a system of equations in terms of p, q, a, b, c, and d. Solve the system for p and q in terms of the components of $\mathbf{A}$.

(c) Find r and s in terms of the components of $\mathbf{A}$.

(d) Under what conditions is it possible to find an inverse of $\mathbf{A}$, and what is that inverse in terms of a, b, c, and d?

(e) Confirm your result from part (d) by multiplying $\mathbf{A}$ by the matrix you found in part (d).

Problem 10.38:

(a) Find the inverse of $\begin{pmatrix} 4 & -3 \\ 5 & 7 \end{pmatrix}$.

(b) Use your answer to part (a) to solve the system of equations $4x - 3y = -2$, $5x + 7y = -3$.

(c) Find the inverse of $\begin{pmatrix} -0.1 & 5 \\ -0.4 & 30 \end{pmatrix}$.

(d) Use your answer to part (c) to solve the system of equations $-0.1t + 5u = -3$, $-0.4t + 30u = 7$.

Problem 10.39: If $\mathbf{A} = \begin{pmatrix} 4 & -3 \\ 2 & 1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} -5 & 3 \\ -8 & 2 \end{pmatrix}$, then find all matrices $\mathbf{M}$ such that $\mathbf{AM} = \mathbf{B}$.

Problem 10.40: Earlier we saw that $\mathbf{AB}$ is not necessarily equal to $\mathbf{BA}$. Must $\mathbf{AA}^{-1}$ equal $\mathbf{A}^{-1}\mathbf{A}$?

Problem 10.41: Suppose $\mathbf{A}$ and $\mathbf{B}$ are invertible. Prove that $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

Problem 10.33: Suppose the product $\mathbf{ABv} = \mathbf{v}$ for all vectors $\mathbf{v}$. What matrix must $\mathbf{AB}$ equal?

Solution for Problem 10.33: In Problem 10.13, we discovered the identity matrix $\mathbf{I}$, which is the unique matrix such that $\mathbf{Iv} = \mathbf{v}$ for all vectors $\mathbf{v}$. If $\mathbf{ABv} = \mathbf{v}$ for all vectors $\mathbf{v}$, then $\mathbf{AB}$ must be this matrix. That is, we must have $\mathbf{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. □

Problem 10.34: If $\mathbf{AB} = \mathbf{I}$, then we say that $\mathbf{A}$ is the **inverse** of $\mathbf{B}$ and that $\mathbf{B}$ is **invertible**. We often denote the inverse of a matrix $\mathbf{B}$ as $\mathbf{B}^{-1}$. Let $\mathbf{M} = \begin{pmatrix} 2 & 3 \\ 6 & 7 \end{pmatrix}$. Find $\mathbf{M}^{-1}$.

Solution for Problem 10.34: We seek the matrix such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 6 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Expanding the product on the left, we have

$$\begin{pmatrix} 2a+6b & 3a+7b \\ 2c+6d & 3c+7d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

From the entries in the first rows, we have $2a + 6b = 1$ and $3a + 7b = 0$. Solving this system gives $a = -\frac{7}{4}$ and $b = \frac{3}{4}$. From the entries in the second rows, we have $2c + 6d = 0$ and $3c + 7d = 1$. Solving this system gives $c = \frac{3}{2}$ and $d = -\frac{1}{2}$. Checking, we find that

$$\begin{pmatrix} -7/4 & 3/4 \\ 3/2 & -1/2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 6 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

so $\mathbf{M}^{-1} = \begin{pmatrix} -7/4 & 3/4 \\ 3/2 & -1/2 \end{pmatrix}$. □

Problem 10.35: Does the matrix $\mathbf{M} = \begin{pmatrix} 3 & 5 \\ -6 & -10 \end{pmatrix}$ have an inverse? If it does, then find it. If it doesn't, explain why not.

Solution for Problem 10.35: Suppose $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is the inverse of $\mathbf{M}$. Then we must have $\mathbf{AM} = \mathbf{I}$, or

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 3 & 5 \\ -6 & -10 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Expanding the product on the left gives

$$\begin{pmatrix} 3a-6b & 5a-10b \\ 3c-6d & 5c-10d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

From the entries in the first rows, we have $3a - 6b = 1$ and $5a - 10b = 0$. The second equation gives $a = 2b$, but substituting this into the first equation gives $6b - 6b = 1$. Uh-oh; this equation has no solution! Therefore, we cannot find a matrix $\mathbf{A}$ such that $\mathbf{AM} = \mathbf{I}$, which means that $\mathbf{M}$ does not have an inverse. □

Why doesn't $\mathbf{M}$ have an inverse? Here's a clue:

Problem 10.36: If $\mathbf{A}$ is invertible, then how are $\det(\mathbf{A})$ and $\det(\mathbf{A}^{-1})$ related?

Solution for Problem 10.36: We have $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, so $\det(\mathbf{A}^{-1}\mathbf{A}) = \det(\mathbf{I}) = (1)(1) - (0)(0) = 1$. We also have $\det(\mathbf{A}^{-1}\mathbf{A}) = \det(\mathbf{A}^{-1}) \cdot \det(\mathbf{A})$, so we must have $\det(\mathbf{A}^{-1}) \cdot \det(\mathbf{A}) = 1$. □

Now we see why the matrix $\mathbf{M} = \begin{pmatrix} 3 & 5 \\ -6 & -10 \end{pmatrix}$ in Problem 10.35 has no inverse. Since

$$\det(\mathbf{M}) = (3)(-10) - (-6)(5) = 0,$$

we know that there can be no matrix $\mathbf{M}^{-1}$ such that $\det(\mathbf{M}^{-1}) \cdot \det(\mathbf{M}) = 1$.

Problem 10.37: Let $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Under what conditions does $\mathbf{A}$ have an inverse? If $\mathbf{A}$ has an inverse, what is that inverse in terms of a, b, c, and d?

Solution for Problem 10.37: Suppose $\mathbf{A}$ is invertible. We let $\mathbf{A}^{-1} = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$, so we must have

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Expanding the product on the left gives

$$\begin{pmatrix} pa + qc & pb + qd \\ ra + sc & rb + sd \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Equating the corresponding entries on the first rows gives the system $pa + qc = 1$, $pb + qd = 0$, and equating the corresponding entries on the second rows gives $ra + sc = 0$, $rb + sd = 1$. If $\begin{vmatrix} a & c \\ b & d \end{vmatrix} \neq 0$, then applying Cramer's rule to solve the system $pa + qc = 1$, $pb + qd = 0$ for p and q gives

$$p = \frac{\begin{vmatrix} 1 & c \\ 0 & d \end{vmatrix}}{\begin{vmatrix} a & c \\ b & d \end{vmatrix}} = \frac{d}{\begin{vmatrix} a & c \\ b & d \end{vmatrix}} \qquad \text{and} \qquad q = \frac{\begin{vmatrix} a & 1 \\ b & 0 \end{vmatrix}}{\begin{vmatrix} a & c \\ b & d \end{vmatrix}} = \frac{-b}{\begin{vmatrix} a & c \\ b & d \end{vmatrix}}.$$

Similarly, if $\begin{vmatrix} a & c \\ b & d \end{vmatrix} \neq 0$, then we can apply Cramer's rule to $ra + sc = 0$, $rb + sd = 1$ to give

$$r = \frac{\begin{vmatrix} 0 & c \\ 1 & d \end{vmatrix}}{\begin{vmatrix} a & c \\ b & d \end{vmatrix}} = \frac{-c}{\begin{vmatrix} a & c \\ b & d \end{vmatrix}} \qquad \text{and} \qquad s = \frac{\begin{vmatrix} a & 0 \\ b & 1 \end{vmatrix}}{\begin{vmatrix} a & c \\ b & d \end{vmatrix}} = \frac{a}{\begin{vmatrix} a & c \\ b & d \end{vmatrix}}.$$

Since $\begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - bc = \det(\mathbf{A})$, the denominator in each of the results for p, q, r, and s equals $\det(\mathbf{A})$, and we have

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Checking, we find that

$$\mathbf{A}^{-1}\mathbf{A} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} da - bc & db - bd \\ -ca + ac & -cb + ad \end{pmatrix} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

as expected.

Our work above shows that we can find an inverse of any 2×2 matrix with a nonzero determinant. We say that such a matrix is **invertible**. Our solution also tells us that if $\mathbf{A}$ is invertible, then its inverse is unique. That is, if $\mathbf{A}$ is invertible, there is only one matrix $\mathbf{A}^{-1}$ such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. □

Important: Let $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. If $\det(\mathbf{A}) = 0$, then $\mathbf{A}$ does not have an inverse. Otherwise, we have

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Problem 10.38:

(a) Find the inverse of $\begin{pmatrix} 4 & -3 \\ 5 & 7 \end{pmatrix}$.

(b) Use your answer to part (a) to solve the system of equations $4x - 3y = -2$, $5x + 7y = -3$.

(c) Find the inverse of $\begin{pmatrix} -0.1 & 5 \\ -0.4 & 30 \end{pmatrix}$.

(d) Use your answer to part (c) to solve the system of equations $-0.1t + 5u = -3$, $-0.4t + 30u = 7$.

Solution for Problem 10.38:

(a) The determinant of the matrix is $(4)(7) - (-3)(5) = 43$, so the inverse of the matrix is

$$\frac{1}{43} \begin{pmatrix} 7 & 3 \\ -5 & 4 \end{pmatrix} = \begin{pmatrix} \frac{7}{43} & \frac{3}{43} \\ -\frac{5}{43} & \frac{4}{43} \end{pmatrix}.$$

(b) We can write the system of equations as $\begin{pmatrix} 4 & -3 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2 \\ -3 \end{pmatrix}$. Multiplying both sides of this equation by the inverse we found in part (a), we have

$$\begin{pmatrix} \frac{7}{43} & \frac{3}{43} \\ -\frac{5}{43} & \frac{4}{43} \end{pmatrix} \begin{pmatrix} 4 & -3 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{7}{43} & \frac{3}{43} \\ -\frac{5}{43} & \frac{4}{43} \end{pmatrix} \begin{pmatrix} -2 \\ -3 \end{pmatrix} = \begin{pmatrix} -\frac{23}{43} \\ -\frac{2}{43} \end{pmatrix}.$$

The product of a matrix and its inverse is the identity, so we have $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -23/43 \\ -2/43 \end{pmatrix}$.

(c) The determinant of the matrix is $(-0.1)(30) - (5)(-0.4) = -1$, so the inverse of the matrix is

$$\frac{1}{-1} \begin{pmatrix} 30 & -5 \\ 0.4 & -0.1 \end{pmatrix} = \begin{pmatrix} -30 & 5 \\ -0.4 & 0.1 \end{pmatrix}.$$

(d) We write the system as $\begin{pmatrix} -0.1 & 5 \\ -0.4 & 30 \end{pmatrix} \begin{pmatrix} t \\ u \end{pmatrix} = \begin{pmatrix} -3 \\ 7 \end{pmatrix}$. Multiplying both sides by the inverse we found in part (c), we have

$$\begin{pmatrix} -30 & 5 \\ -0.4 & 0.1 \end{pmatrix} \begin{pmatrix} -0.1 & 5 \\ -0.4 & 30 \end{pmatrix} \begin{pmatrix} t \\ u \end{pmatrix} = \begin{pmatrix} -30 & 5 \\ -0.4 & 0.1 \end{pmatrix} \begin{pmatrix} -3 \\ 7 \end{pmatrix} = \begin{pmatrix} 125 \\ 1.9 \end{pmatrix}.$$

As before, the product of a matrix and its inverse is the identity, so we have $\begin{pmatrix} t \\ u \end{pmatrix} = \begin{pmatrix} 125 \\ 1.9 \end{pmatrix}$.

□

Problem 10.39: If $\mathbf{A} = \begin{pmatrix} 4 & -3 \\ 2 & 1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} -5 & 3 \\ -8 & 2 \end{pmatrix}$, then find all matrices $\mathbf{M}$ such that $\mathbf{AM} = \mathbf{B}$.

Solution for Problem 10.39: We can't divide by $\mathbf{A}$, but because its determinant is nonzero, we can multiply by its inverse. The determinant of the matrix is 10, so $\mathbf{A}^{-1} = \frac{1}{10} \begin{pmatrix} 1 & 3 \\ -2 & 4 \end{pmatrix} = \begin{pmatrix} 0.1 & 0.3 \\ -0.2 & 0.4 \end{pmatrix}$. What's wrong with this next step:

Bogus Solution: Multiplying both sides of the given equation by $\mathbf{A}^{-1}$ gives us

$$\mathbf{IM} = \mathbf{BA}^{-1}.$$

The problem here is that we started with $\mathbf{AM} = \mathbf{B}$ and multiplied $\mathbf{AM}$ on the left by $\mathbf{A}^{-1}$ to get $\mathbf{IM}$, but we multiplied $\mathbf{B}$ on the right by $\mathbf{A}^{-1}$. While we can go from $t = s$ to $rt = sr$ with real numbers, we can't go from $\mathbf{T} = \mathbf{S}$ to $\mathbf{RT} = \mathbf{SR}$ with matrices, since matrix multiplication is not commutative.

We multiply both sides of the equation $\mathbf{AM} = \mathbf{B}$ on the left by $\mathbf{A}^{-1}$ to produce $\mathbf{A}^{-1}\mathbf{AM} = \mathbf{A}^{-1}\mathbf{B}$. The left side then becomes $\mathbf{IM}$, which equals $\mathbf{M}$, so we have

$$\mathbf{M} = \mathbf{A}^{-1}\mathbf{B} = \begin{pmatrix} 0.1 & 0.3 \\ -0.2 & 0.4 \end{pmatrix}\begin{pmatrix} -5 & 3 \\ -8 & 2 \end{pmatrix} = \begin{pmatrix} -2.9 & 0.9 \\ -2.2 & 0.2 \end{pmatrix}.$$

So, we see that there is a unique solution to the equation $\mathbf{AM} = \mathbf{B}$. Moreover, our solution shows us that whenever $\mathbf{A}$ is invertible, there is a unique solution for $\mathbf{M}$ to the equation $\mathbf{AM} = \mathbf{B}$, which is $\mathbf{M} = \mathbf{A}^{-1}\mathbf{B}$. □

Problem 10.40: Earlier we saw that $\mathbf{AB}$ is not necessarily equal to $\mathbf{BA}$. Must $\mathbf{AA}^{-1}$ equal $\mathbf{A}^{-1}\mathbf{A}$?

Solution for Problem 10.40: We have defined $\mathbf{A}^{-1}$ as the matrix such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, so to check if $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A}$, we need only compute $\mathbf{AA}^{-1}$. Letting $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we have

$$\mathbf{AA}^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\left(\frac{1}{\det(\mathbf{A})}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\right) = \frac{1}{\det(\mathbf{A})}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{\det(\mathbf{A})}\begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} = \mathbf{I}.$$

Therefore, $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. □

We therefore see that if $\mathbf{A}$ is invertible, then the inverse of $\mathbf{A}^{-1}$ is $\mathbf{A}$.

Important: If $\mathbf{A}$ is invertible, then we have $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

Problem 10.41: Suppose $\mathbf{A}$ and $\mathbf{B}$ are invertible. Prove that $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

Solution for Problem 10.41: If $\mathbf{M}$ is the inverse of $\mathbf{AB}$, then we must have $\mathbf{M}(\mathbf{AB}) = \mathbf{I}$. So, we test $\mathbf{M} = \mathbf{B}^{-1}\mathbf{A}^{-1}$, and using the associative property, we find

$$(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) = \mathbf{B}^{-1}(\mathbf{A}^{-1}(\mathbf{AB})) = \mathbf{B}^{-1}((\mathbf{A}^{-1}\mathbf{A})\mathbf{B}) = \mathbf{B}^{-1}\mathbf{IB} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}.$$

Therefore, the inverse of $\mathbf{AB}$ is $\mathbf{B}^{-1}\mathbf{A}^{-1}$. □

Exercises

10.6.1 What is the inverse of $\mathbf{I}$?

10.6.2 For each of the following matrices, find the inverse or show that the matrix does not have an inverse.

(a) $\begin{pmatrix} 6 & 8 \\ -2 & 4 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & -1 \\ -0.9 & 0.5 \end{pmatrix}$ (c) $\begin{pmatrix} -40 & 30 \\ 8 & -6 \end{pmatrix}$

10.6.3 Suppose we are given matrix $\mathbf{A}$ and vector $\mathbf{v}$.

(a) If $\mathbf{A}$ is nonsingular, must there be a vector $\mathbf{w}$ such that $\mathbf{Aw} = \mathbf{v}$?

(b) If $\mathbf{A}$ is singular, can we conclude that there is no vector such that $\mathbf{Aw} = \mathbf{v}$?

10.6.4 All of the entries of $\mathbf{M}$ and $\mathbf{M}^{-1}$ are integers. Find all possible values of $\det(\mathbf{M})$.

10.6.5 In Problem 10.41, we showed that if $\mathbf{A}$ and $\mathbf{B}$ are invertible, then $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$. If $\mathbf{A}$ and $\mathbf{B}$ are invertible, then must we have $(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{B}^{-1} + \mathbf{A}^{-1}$?

10.6.6 In Problem 10.18, we found that multiplying $\begin{pmatrix} x \\ y \end{pmatrix}$ by the matrix $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ corresponds to multiplying the complex number $x + yi$ by $a + bi$. Find the matrix $\mathbf{M}$ such that multiplying $\begin{pmatrix} x \\ y \end{pmatrix}$ by $\mathbf{M}$ corresponds to dividing $x + yi$ by $a + bi$ (assuming a and b are not both 0).

10.7 Summary

A 2×2 matrix represents a linear function that maps vectors in 2 dimensions to vectors in 2 dimensions. This mapping is typically illustrated as a product of the matrix and a vector according to the definition

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

For any matrix $\mathbf{A}$, any scalar c, and any vectors $\mathbf{v}$ and $\mathbf{w}$, we have $\mathbf{A}(c\mathbf{v}) = c(\mathbf{Av})$ and $\mathbf{A}(\mathbf{v} + \mathbf{w}) = \mathbf{Av} + \mathbf{Aw}$.

We add two matrices by adding the corresponding components in the matrices, and define the product of two matrices as follows:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}.$$

Matrix multiplication is *not* commutative, but it is associative and distributive.

Important: If $\mathbf{v}$ and $\mathbf{w}$ are linearly independent and $\mathbf{A}$ and $\mathbf{B}$ are matrices such that $\mathbf{Av} = \mathbf{Bv}$ and $\mathbf{Aw} = \mathbf{Bw}$, then $\mathbf{A} = \mathbf{B}$.

Important: The matrix

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is called the **identity matrix**. Multiplying a vector $\mathbf{v}$ by $\mathbf{I}$ leaves the vector unchanged; in other words, $\mathbf{Iv} = \mathbf{v}$ for all vectors $\mathbf{v}$. We also have $\mathbf{IA} = \mathbf{AI} = \mathbf{A}$ for any 2×2 matrix $\mathbf{A}$.

Important: Any linear function that maps vectors to vectors can be expressed as multiplication by a matrix, where the columns of the matrix are given by the vectors to which $\mathbf{i}$ and $\mathbf{j}$ are mapped by the function.

Definition: The **determinant** of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is the quantity $ad - bc$. We often denote this determinant as $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$, so

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

We also can denote the determinant of a matrix $\mathbf{A}$ as $\det(\mathbf{A})$.

We say a matrix is **singular** if its determinant is 0, and **nonsingular** if its determinant is nonzero. If $\det(\mathbf{A}) \neq 0$, then for any vector $\mathbf{c}$, there is unique vector $\mathbf{v}$ such that $\mathbf{Av} = \mathbf{c}$.

The determinant has the following properties:

- $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a+kc & b+kd \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c+ka & d+kb \end{vmatrix} = \begin{vmatrix} a+kb & b \\ c+kd & d \end{vmatrix} = \begin{vmatrix} a & b+ka \\ c & d+kc \end{vmatrix}$.
- For any 2×2 matrix $\mathbf{A}$ and any constant k, we have $\det(k\mathbf{A}) = k^2 \det(\mathbf{A})$.
- Multiplying either column or either row of a matrix $\mathbf{A}$ by any constant k gives a matrix with determinant $k \det(\mathbf{A})$.
- For any 2×2 matrix $\mathbf{A}$, swapping the rows or columns of $\mathbf{A}$ produces a matrix with determinant $(-1)\det(\mathbf{A})$.
- The determinant of a 2×2 matrix is 0 if and only if the rows of the matrix are linearly dependent.
- We have $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$ for any 2×2 matrices $\mathbf{A}$ and $\mathbf{B}$.
- The absolute value of the determinant $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ equals the area of the parallelogram that has as two sides the vectors $\begin{pmatrix} a \\ c \end{pmatrix}$ and $\begin{pmatrix} b \\ d \end{pmatrix}$ drawn from the origin.

If $\mathbf{BA} = \mathbf{I}$, then we say that $\mathbf{B}$ is the **inverse** of $\mathbf{A}$ and that $\mathbf{A}$ is **invertible**. We typically denote the inverse of $\mathbf{A}$ as $\mathbf{A}^{-1}$, and we have $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

Important: Let $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. If $\det(\mathbf{A}) = 0$, then $\mathbf{A}$ does not have an inverse. Otherwise, we have

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

REVIEW PROBLEMS

10.42 Find constants a, b, and c such that $\begin{pmatrix} a & 4 \\ 7 & 2 \end{pmatrix} + c \begin{pmatrix} 6 & b \\ -3 & -1 \end{pmatrix} = \begin{pmatrix} 32 & 9 \\ -8 & -3 \end{pmatrix}$.

10.43 Let $\mathbf{M} = \begin{pmatrix} 2 & 3 \\ -7 & 0 \end{pmatrix}$ and $\mathbf{N} = \begin{pmatrix} -3 & -1 \\ 5 & 2 \end{pmatrix}$. Find $\mathbf{MN}$ and $\mathbf{NM}$.

10.44 Show that the rows of a matrix are linearly dependent if and only if the columns of the matrix are linearly dependent.

10.45 Is there a matrix $\mathbf{F}$ such that for all matrices $\mathbf{A}$, the product $\mathbf{FA}$ is the matrix formed by swapping the columns of $\mathbf{A}$? If such a matrix exists, what is $\mathbf{AF}$ for that matrix?

10.46 Suppose $\mathbf{Av} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$ and $\mathbf{Aw} = \begin{pmatrix} -7 \\ 4 \end{pmatrix}$. Find $\mathbf{A}(3\mathbf{v} - 4\mathbf{w})$.

10.47

(a) Find two matrices besides $\mathbf{0}$ and the identity such that $\mathbf{A}^2 = \mathbf{A}$.

(b) Find a matrix besides $\mathbf{0}$ and the identity such that $\mathbf{A}^3 = \mathbf{A}$ and $\mathbf{A}^3 \neq \mathbf{A}^2$.

10.48

(a) Find a matrix $\mathbf{A}$ such that $\mathbf{Av}$ is the reflection of $\mathbf{v}$ over the y-axis.

(b) Find a matrix $\mathbf{B}$ such that $\mathbf{Bv}$ is the reflection of $\mathbf{v}$ over the graph of $x = y$.

10.49 What is the matrix $\mathbf{R}$ such that $\mathbf{Rv}$ is a *clockwise* rotation of $\mathbf{v}$ by an angle θ for all vectors $\mathbf{v}$?

10.50 Let $\mathbf{T} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Compute $\mathbf{T}^2$ and $\mathbf{T}^3$ and confirm that these equal the rotation matrices for angles of 2θ and 3θ, respectively.

10.51 Find $\mathbf{M}^{10}$ if $\mathbf{M} = \begin{pmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & \sqrt{2} \end{pmatrix}$.

10.52 Point P is reflected over the y-axis and the result is rotated 60° counterclockwise about the origin. The point that results is $(-2 - \sqrt{3}, 1 - 2\sqrt{3})$. What are the coordinates of point P?

10.53 Show that $f(\mathbf{0}) = \mathbf{0}$ for any linear function f that maps vectors to vectors.

10.54 Evaluate the following:

(a) $\begin{vmatrix} 7 & -4 \\ -2 & 7 \end{vmatrix}$ (b) $\begin{vmatrix} 0.27 & -175 \\ -0.15 & 95 \end{vmatrix}$

10.55

(a) Must the product of two singular matrices be singular?

(b) Must the product of two nonsingular matrices be nonsingular?

(c) Is the product of a singular matrix and nonsingular matrix always singular, always nonsingular, or sometimes singular and sometimes nonsingular?

10.56 Suppose we form the first row of $\mathbf{C}$ by adding k times the second row of $\mathbf{A}$ to the first row of $\mathbf{A}$, and we form the second row of $\mathbf{C}$ by adding j times the first row of $\mathbf{A}$ to the second row of $\mathbf{A}$. Must the determinants of $\mathbf{A}$ and $\mathbf{C}$ be the same?

10.57 Explain geometrically why swapping the columns of $\mathbf{A}$ produces a matrix with determinant $(-1)\det(\mathbf{A})$.

10.58 For each of the following matrices, find the inverse, or show that the matrix does not have an inverse.

(a) $\begin{pmatrix} -2 & 3 \\ -2 & 8 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & -3 \\ -2 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 6 & 12 \\ -2 & -4 \end{pmatrix}$

10.59 Suppose that $\mathbf{A}^{-1} = \mathbf{A}$. Find all possible values of $\det(\mathbf{A})$.

10.60 The **transpose** of a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$. We denote the transpose of $\mathbf{A}$ as $\mathbf{A}^T$.

(a) Show that $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$.

(b) If $\mathbf{A}$ is invertible, must $\mathbf{A}^T$ be invertible?

(c) If $\mathbf{A}^T$ is invertible, must $(\mathbf{A}^T)^{-1}$ equal $(\mathbf{A}^{-1})^T$?

(d) Let $\mathbf{R} = \begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix}$ and $\mathbf{S} = \begin{pmatrix} -1 & 5 \\ -6 & 3 \end{pmatrix}$. Find $\mathbf{RS}$, $(\mathbf{RS})^T$, $\mathbf{R}^T\mathbf{S}^T$, and $\mathbf{S}^T\mathbf{R}^T$. If you notice anything interesting in your results, try to prove that what you noticed holds for any matrices $\mathbf{R}$ and $\mathbf{S}$.

Challenge Problems

10.61 Find two nonzero matrices $\mathbf{A}$ and $\mathbf{B}$ such that neither matrix is the identity, $\mathbf{A} \neq \mathbf{B}$, and $\mathbf{AB} = \mathbf{BA} \neq \mathbf{0}$. As an extra challenge, find a pair such that no entry in either matrix is 0. **Hints:** 118

10.62 Suppose $\mathbf{A}^6 = \mathbf{A}$. Find all possible values of $\det(\mathbf{A})$, assuming all the entries of $\mathbf{A}$ are real. What if we allow the entries of $\mathbf{A}$ to be nonreal?

10.63 Find a matrix such that $\mathbf{A}^4 = \mathbf{A}$, but $\mathbf{A}^2 \neq \mathbf{A}$. **Hints:** 230, 202

10.64 Let $\mathbf{A}$ and $\mathbf{B}$ be two 2×2 matrices. Prove that $\det(\mathbf{A}+\mathbf{B}) + \det(\mathbf{A}-\mathbf{B}) = 2(\det(\mathbf{A}) + \det(\mathbf{B}))$.

10.65 Explain why the line through the points (x_1, y_1) and (x_2, y_2) is the graph of the equation $\begin{vmatrix} x - x_1 & y - y_1 \\ x - x_2 & y - y_2 \end{vmatrix} = 0$.

10.66★ Let $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$.

(a) Prove that $\mathbf{A}^n = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}$, where F_n denotes the n^{th} Fibonacci number.

(b) Prove that $F_{n+1}F_{n-1} - F_n^2 = (-1)^n$ for all $n \geq 0$. **Hints:** 136

10.67 Describe all (x, y) such that there exists a value of θ for which $x\cos\theta - y\sin\theta = 3$ and $x\sin\theta + y\cos\theta = 2$. **Hints:** 115

10.68

(a) Show that reflection over any line through the origin is a linear transformation.

(b) Find the reflection of $(4, -6)$ over the line $y = 2x$.

(c) Let f be a function that maps a point (x, y) to the image of the reflection of (x, y) over the line $x + y = 6$. Is f a linear function?

10.69 Let $\mathbf{A} = \begin{pmatrix} -2 & 6 \\ 2 & -1 \end{pmatrix}$.

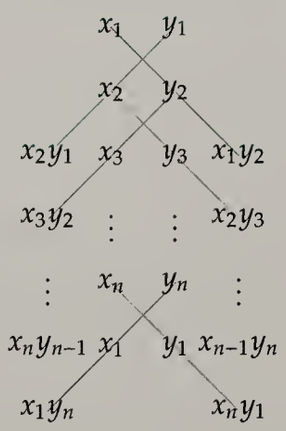
(a) Find all nonzero vectors $\mathbf{v}$ such that there exists a constant λ for which $\mathbf{Av} = \lambda\mathbf{v}$.

(b) For each vector $\mathbf{v}$ and corresponding λ you find in part (a), find $\det(\mathbf{A} - \lambda\mathbf{I})$. Explain your results. **Hints:** 213

10.70 Suppose the columns of $\mathbf{A}$ are orthogonal. Find both a geometric and an algebraic proof that $|\det(\mathbf{A})|$ equals the product of the norms of the columns.

10.71★ Suppose that the points (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) are the vertices of a convex polygon, in counterclockwise order. We can find the area of this polygon with the following process:

1. List the vertices in order, as shown in the middle two columns at right, and include (x_1, y_1) a second time at the end of the list.
2. Take pairwise diagonal products, as shown in the first and fourth columns at right.
3. Sum the first column and sum the fourth column. The area of the polygon then is half the difference of the fourth column sum and the first column sum.



This fact is sometimes sometimes referred to **Shoelace Theorem**.

(a) Prove the Shoelace Theorem for $n = 3$.

(b) Prove the Shoelace Theorem for $n = 4$. **Hints:** 18

(c) Prove the Shoelace Theorem for any integer n with $n \geq 3$. **Hints:** 287

10.72 Back on page 367, we made the claim that the parallelogram with sides $\mathbf{Ar}$ and $\mathbf{As}$ has signed area $\det(\mathbf{A})$ times the signed area of the parallelogram with sides $\mathbf{r}$ and $\mathbf{s}$, where the orientation is determined by traveling first from the origin along $\mathbf{r}$. We proved this statement when $\mathbf{r}$ and s are $\mathbf{i}$ and $\mathbf{j}$. In this problem, we'll prove it for any linearly independent pair of vectors $\mathbf{r}$ and s.

(a) Why must there exist a matrix $\mathbf{B}$ such that $\mathbf{Bi} = \mathbf{r}$ and $\mathbf{Bj} = \mathbf{s}$?

(b) In terms of $\mathbf{A}$ and $\mathbf{B}$, what is the signed area of the parallelogram with sides $\mathbf{r}$ and $\mathbf{s}$?

(c) In terms of $\mathbf{A}$ and $\mathbf{B}$, what is the signed area of the parallelogram with sides $\mathbf{Ar}$ and $\mathbf{As}$? Note that you cannot simply multiply your answer from part (b) by $\det(\mathbf{A})$; we are trying to prove such a multiplication works!

(d) Complete the proof by showing that the parallelogram with sides $\mathbf{Ar}$ and $\mathbf{As}$ has signed area $\det(\mathbf{A})$ times the signed area of the parallelogram with sides $\mathbf{r}$ and s.

10.73 For any vector $\mathbf{v}$, the product $\mathbf{Tv}$ is the reflection of $\mathbf{v}$ over the line $y = 3x$. Find $\det(\mathbf{T})$. **Hints:** 144

10.74 Find all matrices $\mathbf{A}$ such that $\mathbf{A}^2 = \mathbf{I}$. **Hints:** 75, 116

10.75 Is it true that $\mathbf{A}^2 - \mathbf{B}^2 = (\mathbf{A} - \mathbf{B})(\mathbf{A} + \mathbf{B})$ for all matrices $\mathbf{A}$ and $\mathbf{B}$?

10.76★ Let $\mathbf{A}$ and $\mathbf{B}$ be matrices such that $\mathbf{A} + \mathbf{B} = \mathbf{AB}$. Show that $\mathbf{AB} = \mathbf{BA}$. **Hints:** 222, 277, 54

10.77★ Let $\mathbf{A}$ be a 2×2 matrix such that $\mathbf{A}\begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}4\\5\end{pmatrix}$ and $\mathbf{A}^2 - \mathbf{A} + 5\mathbf{I} = \mathbf{0}$. Find $\mathbf{A}$. **Hints:** 119

10.78

(a) If $\mathbf{A}^2 = \mathbf{0}$, then must $\mathbf{A} = \mathbf{0}$?

(b)★ If $\mathbf{A}^3 = \mathbf{0}$, then must $\mathbf{A}^2 = \mathbf{0}$? **Hints:** 221

Every new body of discovery is mathematical in form, because there is no other guidance we can have. – Charles Darwin

CHAPTER 11

Vectors and Matrices in Three Dimensions, Part 1

11.1 Vectors in Three Dimensions

Vectors in three dimensions have much in common with vectors in two dimensions. Just as we denote a vector in two dimensions with an ordered pair $\begin{pmatrix} x \\ y \end{pmatrix}$, we denote a vector in three dimensions with an ordered triple $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$.

We define vector addition and multiplication by a scalar for vectors in three dimensions in essentially the same way we did in two dimensions:

- **Vector addition**. We add vectors by adding the corresponding components:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a+x \\ b+y \\ c+z \end{pmatrix}.$$

- **Multiplication of a vector by a scalar**. We multiply a vector by a scalar by multiplying each component of the vector by the scalar:

$$k\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} ka \\ kb \\ kc \end{pmatrix}.$$

We also define the dot product of two vectors in three dimensions much as we defined it for vectors in two

dimensions:

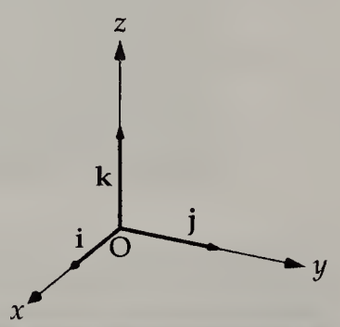
$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = v_1 \cdot w_1 + v_2 \cdot w_2 + v_3 \cdot w_3.$$

Just as we can represent the vector $\begin{pmatrix} x \\ y \end{pmatrix}$ as an arrow from the origin in the Cartesian plane to the point (x, y), we can represent the vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ as an arrow from the origin to the point (x, y, z) in space. With this in mind, we can extend the concept of the **norm** of a vector to vectors in three dimensions. We define the norm of $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ as the distance between (x, y, z) and the origin. We showed on page 178 that the distance between (x_1, y_1, z_1) and (x_2, y_2, z_2) in space is $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$. Therefore, the distance from (x, y, z) to the origin is $\sqrt{x^2 + y^2 + z^2}$. As in two dimensions, we denote the norm of $\mathbf{v}$ as $\|\mathbf{v}\|$, so we have

$$\left\| \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right\| = \sqrt{x^2 + y^2 + z^2}.$$

Also as in two dimensions, the norm of a vector is sometimes referred to as the vector's **length** or **magnitude**.

A vector with magnitude 1 is called a **unit vector**. The vector $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ is also written $a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$, where $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are unit vectors from the origin in the direction of the positive x-axis, y-axis, and z-axis, respectively. Just as with vectors in two dimensions, a nonzero vector in three dimensions can be written as the product of its magnitude and a unit vector that is sometimes referred to as the **direction** of the vector.



Problems

Problem 11.1: Let $\mathbf{u} = \begin{pmatrix} 1 \\ 3 \\ -5 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 0 \\ -4 \\ 5 \end{pmatrix}$, and $\mathbf{w} = \begin{pmatrix} x \\ 2 \\ y \end{pmatrix}$.

(a) Find $\mathbf{u} + \mathbf{v}$.

(b) Find $3\mathbf{u} - 5\mathbf{v}$.

(c) Find x and y if $2\mathbf{w} - \mathbf{u} = t\mathbf{v}$ for some constant t.

Problem 11.2: Evaluate $\begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 3 \\ 5 \end{pmatrix}$ and $(-5\mathbf{i} + 3\mathbf{j} - 2\mathbf{k}) \cdot (5\mathbf{i} - 4\mathbf{j} + 0\mathbf{k})$.

Problem 11.3: Show that for any scalar c and any vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, we have:

(a) $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$

(b) $\mathbf{v} \cdot (c\mathbf{w}) = c(\mathbf{v} \cdot \mathbf{w})$

(c) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$

(d) $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$

Problem 11.4: Show that if $\mathbf{v}$ and $\mathbf{w}$ are nonzero vectors in three dimensions, then $\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos\theta$, where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$ when these vectors are drawn from the origin in space.

We start with a little practice adding and subtracting vectors in three dimensions.

Problem 11.1: Let $\mathbf{u} = \begin{pmatrix} 1 \\ 3 \\ -5 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 0 \\ -4 \\ 5 \end{pmatrix}$, and $\mathbf{w} = \begin{pmatrix} x \\ 2 \\ y \end{pmatrix}$.

(a) Find $\mathbf{u} + \mathbf{v}$.

(b) Find $3\mathbf{u} - 5\mathbf{v}$.

(c) Find x and y if $2\mathbf{w} - \mathbf{u} = t\mathbf{v}$ for some constant t.

Solution for Problem 11.1:

(a) We have $\begin{pmatrix} 1 \\ 3 \\ -5 \end{pmatrix} + \begin{pmatrix} 0 \\ -4 \\ 5 \end{pmatrix} = \begin{pmatrix} 1+0 \\ 3-4 \\ -5+5 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$.

(b) We have $3\begin{pmatrix} 1 \\ 3 \\ -5 \end{pmatrix} - 5\begin{pmatrix} 0 \\ -4 \\ 5 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \\ -15 \end{pmatrix} - \begin{pmatrix} 0 \\ -20 \\ 25 \end{pmatrix} = \begin{pmatrix} 3 \\ 29 \\ -40 \end{pmatrix}$.

(c) Since $2\begin{pmatrix} x \\ 2 \\ y \end{pmatrix} - \begin{pmatrix} 1 \\ 3 \\ -5 \end{pmatrix} = \begin{pmatrix} 2x-1 \\ 1 \\ 2y+5 \end{pmatrix}$ and $t\begin{pmatrix} 0 \\ -4 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ -4t \\ 5t \end{pmatrix}$, we must have $2x + 1 = 0$, $1 = -4t$, and $2y + 5 = 5t$. From the first equation, we have $x = -\frac{1}{2}$. From the second equation, we have $t = -\frac{1}{4}$, and substituting this value of t into the last equation gives $y = -\frac{25}{8}$.

□

Problem 11.2: Evaluate $\begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 3 \\ 5 \end{pmatrix}$ and $(-5\mathbf{i} + 3\mathbf{j} - 2\mathbf{k}) \cdot (5\mathbf{i} - 4\mathbf{j} + 0\mathbf{k})$.

Solution for Problem 11.2: Applying the definition for the dot product from the introduction, we have

$$\begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 3 \\ 5 \end{pmatrix} = (2)(-1) + (-3)(3) + (4)(5) = -2 - 9 + 20 = 9$$

and

$$(-5\mathbf{i} + 3\mathbf{j} - 2\mathbf{k}) \cdot (5\mathbf{i} - 4\mathbf{j} + 0\mathbf{k}) = (-5)(5) + (3)(-4) + (-2)(0) = -37.$$

□

Problem 11.3: Show that for any scalar c and any vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, we have:

(a) $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$

(b) $\mathbf{v} \cdot (c\mathbf{w}) = c(\mathbf{v} \cdot \mathbf{w})$

(c) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$

(d) $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$

Solution for Problem 11.3: In each of the following parts, we let $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$, $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$, and $\mathbf{w} = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k}$.

(a) We have $\mathbf{v} \cdot \mathbf{w} = v_1w_1 + v_2w_2 + v_3w_3 = w_1v_1 + w_2v_2 + w_3v_3 = \mathbf{w} \cdot \mathbf{v}$.

(b) We have

$$\mathbf{v} \cdot (c\mathbf{w}) = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \cdot \begin{pmatrix} cw_1 \\ cw_2 \\ cw_3 \end{pmatrix} = (v_1)(cw_1) + (v_2)(cw_2) + (v_3)(cw_3) = c(v_1w_1 + v_2w_2 + v_3w_3) = c(\mathbf{v} \cdot \mathbf{w}).$$

(c) We have

$$\begin{aligned} \mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) &= u_1(v_1 + w_1) + u_2(v_2 + w_2) + u_3(v_3 + w_3) \\ &= (u_1v_1 + u_2v_2 + u_3v_3) + (u_1w_1 + u_2w_2 + u_3w_3) \\ &= \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}. \end{aligned}$$

(d) We have $\mathbf{v} \cdot \mathbf{v} = v_1^2 + v_2^2 + v_3^2 = \|\mathbf{v}\|^2$.

□

So, we see that the basic properties of the dot product that we showed for two dimensions also hold for three dimensions.

Important: For any scalar c and any vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, we have:

(a) $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$

(b) $\mathbf{v} \cdot (c\mathbf{w}) = c(\mathbf{v} \cdot \mathbf{w})$

(c) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$

(d) $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$

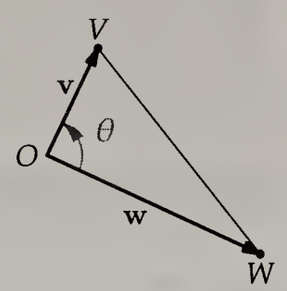
Combining properties (a), (b), and (c) above tells us that the dot product of vectors in three dimensions is **bilinear**, just like the dot product of vectors in two dimensions. By this, we mean that for any vector $\mathbf{v}$, the functions $f(\mathbf{w}) = \mathbf{v} \cdot \mathbf{w}$ and $f(\mathbf{w}) = \mathbf{w} \cdot \mathbf{v}$ are both linear functions.

We also discovered a nice geometric interpretation of the dot product of vectors in two dimensions. Let's see if this interpretation holds for vectors in three dimensions, as well.

Problem 11.4: Show that if $\mathbf{v}$ and $\mathbf{w}$ are nonzero vectors in three dimensions, then $\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \, \|\mathbf{w}\| \cos\theta$, where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$ when these vectors are drawn from the origin in space.

Solution for Problem 11.4: If we look back to our proof for vectors in two dimensions in Problem 9.9 on page 309, we have a pleasant surprise: the same proof holds in three dimensions!

At right, we have points V and W, where $\vec{V} = \mathbf{v}$ and $\vec{W} = \mathbf{w}$, and we let the origin be O. We therefore have $OV = \|\mathbf{v}\|$, $OW = \|\mathbf{w}\|$, and $WV = \left\|\vec{WV}\right\| = \|\mathbf{v} - \mathbf{w}\|$.



The Law of Cosines gives us $WV^2 = OV^2 + OW^2 - 2(OV)(OW)\cos\theta$, or

$$\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\|\mathbf{v}\| \, \|\mathbf{w}\| \cos\theta.$$

Since $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$, $\|\mathbf{w}\|^2 = \mathbf{w} \cdot \mathbf{w}$, and $\|\mathbf{v} - \mathbf{w}\|^2 = (\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w})$, we can write the Law of Cosines equation above as

$$(\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w}) = \mathbf{v} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w} - 2\|\mathbf{v}\| \, \|\mathbf{w}\| \cos\theta.$$

Expanding the dot product on the left gives

$$\mathbf{v}\cdot\mathbf{v} - 2\mathbf{v}\cdot\mathbf{w} + \mathbf{w}\cdot\mathbf{w} = \mathbf{v}\cdot\mathbf{v} + \mathbf{w}\cdot\mathbf{w} - 2\|\mathbf{v}\|\,\|\mathbf{w}\|\cos\theta$$

and simplifying this equation gives the familiar

$$\mathbf{v}\cdot\mathbf{w} = \|\mathbf{v}\|\,\|\mathbf{w}\|\cos\theta.$$

□

If you look back to our proof for vectors in two dimensions on page 309, you'll see that it looks *very* similar to our proof above!

Important: If θ is the angle between $\mathbf{v}$ and $\mathbf{w}$, then

$$\mathbf{v}\cdot\mathbf{w} = \|\mathbf{v}\|\,\|\mathbf{w}\|\cos\theta.$$

As in two dimensions, we can use this relationship to determine the angle between two vectors, and if two vectors have dot product 0, then we say they are **orthogonal**. If orthogonal vectors are nonzero, then their geometric representations are perpendicular, since $\mathbf{v}\cdot\mathbf{w} = 0$ for nonzero $\mathbf{v}$ and $\mathbf{w}$ if and only if the cosine of the angle between them is 0.

Exercises

11.1.1 Let $\mathbf{u} = 4\mathbf{i} + 2\mathbf{j} - \mathbf{k}$ and $\mathbf{v} = -\mathbf{i} + 5\mathbf{j} + 2\mathbf{k}$.

(a) Find $\mathbf{w}$ if $\mathbf{u} - 3\mathbf{w} = 2\mathbf{v}$.

(b) Do there exist nonzero scalars a and b such that $a\mathbf{u} + b\mathbf{v} = \mathbf{0}$?

(c) Does there exist a vector $\mathbf{r}$ such that there is no pair of scalars (x, y) for which $x\mathbf{u} + y\mathbf{v} = \mathbf{r}$?

11.1.2 For what values of x and y is $\begin{pmatrix} 1 \\ x \\ y \end{pmatrix}$ orthogonal to both $\begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix}$ and $\begin{pmatrix} -4 \\ 1 \\ 2 \end{pmatrix}$?

11.1.3 Find the angle between $\begin{pmatrix} 3 \\ -4 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 4 \\ 7 \end{pmatrix}$ to the nearest degree.

11.1.4 Suppose the point (x, y, z) in space has spherical coordinates (ρ, θ, ϕ), where $\rho \geq 0$. If $\mathbf{u} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, then express ϕ in terms of $\mathbf{u}$, $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$.

11.1.5 Let $\mathbf{v} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ and $\mathbf{w} = \mathbf{i} - 2\mathbf{j} + 4\mathbf{k}$.

(a) Describe the graph of the points (x, y, z) such that $\|\mathbf{v}\| = 4$.

(b) Describe the graph of the points (x, y, z) such that $\|\mathbf{v} - \mathbf{w}\| = 6$.

(c)★ Describe the graph of the points (x, y, z) such that $2\mathbf{v}\cdot\mathbf{w} = \|\mathbf{v}\|\,\|\mathbf{w}\|$

11.2 3×3 Matrices

Just as we used a 2×2 matrix to help represent a system of two linear equations in two variables, we can use a 3×3 matrix to help represent a system of three linear equations in three variables. Specifically, we can represent

the system

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= b_1, \\ a_{21}x + a_{22}y + a_{23}z &= b_2, \\ a_{31}x + a_{32}y + a_{33}z &= b_3, \end{aligned}$$

as

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

The product of the 3×3 matrix and the vector on the left is defined as

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y + a_{13}z \\ a_{21}x + a_{22}y + a_{23}z \\ a_{31}x + a_{32}y + a_{33}z \end{pmatrix}.$$

Just as a 2×2 matrix represents a function that maps vectors in two dimensions to vectors in two dimensions, a 3×3 matrix represents a function that maps vectors in three dimensions to vectors in three dimensions. We say that such a function has domain $\mathbb{R}^3$.

Also as with 2×2 matrices, we typically use a capital letter to denote a 3×3 matrix, and often use the lowercase of the letter together with subscripts to denote individual entries of the matrix. The first subscript refers to the row of the entry and the second refers to the column. So, for example, s_{13} refers to the entry of matrix $\mathbf{S}$ in the first row of the third column.

Adding two 3×3 matrices and multiplying a 3×3 matrix by a scalar work essentially the same way they do for 2×2 matrices:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} \\ a_{31} + b_{31} & a_{32} + b_{32} & a_{33} + b_{33} \end{pmatrix}$$

$$k \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} ka_{11} & ka_{12} & ka_{13} \\ ka_{21} & ka_{22} & ka_{23} \\ ka_{31} & ka_{32} & ka_{33} \end{pmatrix}.$$

Problems

Problem 11.5: Compute the following products:

(a) $\begin{pmatrix} 1 & 3 & -4 \\ -3 & 0 & 5 \\ -6 & 1 & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 5 \end{pmatrix}$

(b) $\begin{pmatrix} 0 & -5 & 2 \\ -1 & 1 & -3 \\ 5 & -2 & -1 \end{pmatrix} \begin{pmatrix} -3 \\ 2 \\ 6 \end{pmatrix}$

Problem 11.6: Show that $\mathbf{A}(k\mathbf{v}) = k(\mathbf{Av})$ and $\mathbf{A}(\mathbf{v} + \mathbf{w}) = \mathbf{Av} + \mathbf{Aw}$ for any 3×3 matrix $\mathbf{A}$, any scalar k, and any vectors $\mathbf{v}$ and $\mathbf{w}$ in three dimensions.

Problem 11.7: In this problem, we find the product of two 3×3 matrices. Consider the system of equations

$$\begin{aligned} 3x + 5y - 2z &= p, \\ y - 3z &= q, \\ -2x + 4y + z &= r. \end{aligned}$$

Suppose we also have $x = 3t - 4u - v$, $y = -t + v$, and $z = 7t - 3u - v$.

(a) Find matrix $\mathbf{A}$ such that the original system of equations can be expressed as $\mathbf{A}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} p \\ q \\ r \end{pmatrix}$.

(b) Find matrix $\mathbf{B}$ such that $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{B}\begin{pmatrix} t \\ u \\ v \end{pmatrix}$.

(c) Suppose we wish to define the product $\mathbf{AB}$ such that $(\mathbf{AB})\begin{pmatrix} t \\ u \\ v \end{pmatrix} = \mathbf{A}\left(\mathbf{B}\begin{pmatrix} t \\ u \\ v \end{pmatrix}\right)$ for all $\begin{pmatrix} t \\ u \\ v \end{pmatrix}$. Find the matrix that equals $\mathbf{AB}$.

Problem 11.8: Find the product

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}.$$

Problem 11.9: Find the following products:

(a) $\begin{pmatrix} 2 & -3 & 0 \\ -1 & -2 & 5 \\ -2 & -5 & 1 \end{pmatrix}\begin{pmatrix} 4 & -1 & 2 \\ -1 & 1 & -2 \\ 5 & 0 & -2 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & -1 & 4 \\ 0 & 1 & 7 \\ 2 & -1 & 3 \end{pmatrix}\begin{pmatrix} -1 & 4 & 3 \\ 3 & 5 & -1 \\ 1 & 1 & 5 \end{pmatrix}$

Problem 11.10: If $\mathbf{A}$ and $\mathbf{B}$ are 3×3 matrices, must we have $\mathbf{AB} = \mathbf{BA}$?

Problem 11.11: If $\mathbf{A}$ and $\mathbf{B}$ are 3×3 matrices such that $\mathbf{Ai} = \mathbf{Bi}$, $\mathbf{Aj} = \mathbf{Bj}$, and $\mathbf{Ak} = \mathbf{Bk}$, then must we have $\mathbf{A} = \mathbf{B}$?

Problem 11.12:

(a) Find a 3×3 matrix $\mathbf{I}$ such that $\mathbf{IA} = \mathbf{A}$ for any 3×3 matrix $\mathbf{A}$.

(b) Do we also have $\mathbf{AI} = \mathbf{A}$ for this matrix $\mathbf{I}$?

(c) Is the matrix you found in part (a) the only matrix $\mathbf{I}$ such that $\mathbf{IA} = \mathbf{A}$ for all $\mathbf{A}$? Is it the only matrix such that $\mathbf{AI} = \mathbf{A}$ for all $\mathbf{A}$?

Problem 11.13:

(a) Find the matrix $\mathbf{P}$ such that for every $\mathbf{A}$, the product $\mathbf{PA}$ results in a matrix whose third row is the same as that of $\mathbf{A}$, but the first two rows are swapped. (That is, the second row of the product is the first row of $\mathbf{A}$ and vice-versa.)

(b) Find a matrix $\mathbf{T}$ such that $\mathbf{T} \neq \mathbf{I}$ and $\mathbf{T}^3 = \mathbf{I}$.

Problem 11.5: Compute the following products:

(a) $\begin{pmatrix} 1 & 3 & -4 \\ -3 & 0 & 5 \\ -6 & 1 & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 5 \end{pmatrix}$

(b) $\begin{pmatrix} 0 & -5 & 2 \\ -1 & 1 & -3 \\ 5 & -2 & -1 \end{pmatrix} \begin{pmatrix} -3 \\ 2 \\ 6 \end{pmatrix}$

Solution for Problem 11.5:

(a) Applying the definition of multiplication of a vector by a matrix, we have

$$\begin{pmatrix} 1 & 3 & -4 \\ -3 & 0 & 5 \\ -6 & 1 & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} (1)(0) + (3)(3) + (-4)(5) \\ (-3)(0) + (0)(3) + (5)(5) \\ (-6)(0) + (1)(3) + (-2)(5) \end{pmatrix} = \begin{pmatrix} -11 \\ 25 \\ -7 \end{pmatrix}.$$

(b) We have

$$\begin{pmatrix} 0 & -5 & 2 \\ -1 & 1 & -3 \\ 5 & -2 & -1 \end{pmatrix} \begin{pmatrix} -3 \\ 2 \\ 6 \end{pmatrix} = \begin{pmatrix} (0)(-3) + (-5)(2) + (2)(6) \\ (-1)(-3) + (1)(2) + (-3)(6) \\ (5)(-3) + (-2)(2) + (-1)(6) \end{pmatrix} = \begin{pmatrix} 2 \\ -13 \\ -25 \end{pmatrix}.$$

□

Problem 11.6: Show that $\mathbf{A}(k\mathbf{v}) = k(\mathbf{Av})$ and $\mathbf{A}(\mathbf{v}+\mathbf{w}) = \mathbf{Av} + \mathbf{Aw}$ for any 3×3 matrix $\mathbf{A}$, any scalar k, and any vectors $\mathbf{v}$ and $\mathbf{w}$ in three dimensions.

Solution for Problem 11.6: We have

$$\mathbf{A}(k\mathbf{v}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} kv_1 \\ kv_2 \\ kv_3 \end{pmatrix} = \begin{pmatrix} a_{11}kv_1 + a_{12}kv_2 + a_{13}kv_3 \\ a_{21}kv_1 + a_{22}kv_2 + a_{23}kv_3 \\ a_{31}kv_1 + a_{32}kv_2 + a_{33}kv_3 \end{pmatrix} = k \begin{pmatrix} a_{11}v_1 + a_{12}v_2 + a_{13}v_3 \\ a_{21}v_1 + a_{22}v_2 + a_{23}v_3 \\ a_{31}v_1 + a_{32}v_2 + a_{33}v_3 \end{pmatrix} = k(\mathbf{Av})$$

and

$$\begin{aligned} \mathbf{A}(\mathbf{v}+\mathbf{w}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{pmatrix} &= \begin{pmatrix} a_{11}(v_1 + w_1) + a_{12}(v_2 + w_2) + a_{13}(v_3 + w_3) \\ a_{21}(v_1 + w_1) + a_{22}(v_2 + w_2) + a_{23}(v_3 + w_3) \\ a_{31}(v_1 + w_1) + a_{32}(v_2 + w_2) + a_{33}(v_3 + w_3) \end{pmatrix} \\ &= \begin{pmatrix} a_{11}v_1 + a_{12}v_2 + a_{13}v_3 \\ a_{21}v_1 + a_{22}v_2 + a_{23}v_3 \\ a_{31}v_1 + a_{32}v_2 + a_{33}v_3 \end{pmatrix} + \begin{pmatrix} a_{11}w_1 + a_{12}w_2 + a_{13}w_3 \\ a_{21}w_1 + a_{22}w_2 + a_{23}w_3 \\ a_{31}w_1 + a_{32}w_2 + a_{33}w_3 \end{pmatrix} \\ &= \mathbf{Av} + \mathbf{Aw}. \end{aligned}$$

□

Important: For any 3×3 matrix $\mathbf{A}$, any scalar k, and any vectors $\mathbf{v}$ and $\mathbf{w}$ in three dimensions, we have $\mathbf{A}(k\mathbf{v}) = k\mathbf{Av}$ and $\mathbf{A}(\mathbf{v}+\mathbf{w}) = \mathbf{Av} + \mathbf{Aw}$.

So, we see that the function that a 3×3 matrix represents is a linear function.

Problem 11.7: In this problem, we find the product of two 3×3 matrices. Consider the system of equations

$$\begin{aligned} 3x + 5y - 2z &= p, \\ y - 3z &= q, \\ -2x + 4y + z &= r. \end{aligned}$$

Suppose we also have $x = 3t - 4u - v$, $y = -t + v$, and $z = 7t - 3u - v$.

(a) Find matrix $\mathbf{A}$ such that the original system of equations can be expressed as $\mathbf{A}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} p \\ q \\ r \end{pmatrix}$.

(b) Find matrix $\mathbf{B}$ such that $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{B}\begin{pmatrix} t \\ u \\ v \end{pmatrix}$.

(c) Suppose we wish to define the product $\mathbf{AB}$ such that $(\mathbf{AB})\begin{pmatrix} t \\ u \\ v \end{pmatrix} = \mathbf{A}\left(\mathbf{B}\begin{pmatrix} t \\ u \\ v \end{pmatrix}\right)$ for all $\begin{pmatrix} t \\ u \\ v \end{pmatrix}$. Find the matrix that equals $\mathbf{AB}$.

Solution for Problem 11.7:

(a) Following the definition of a 3×3 matrix in the introduction, we can write the initial system of equations as

$$\begin{pmatrix} 3 & 5 & -2 \\ 0 & 1 & -3 \\ -2 & 4 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} p \\ q \\ r \end{pmatrix}.$$

(b) Again using our definition of a 3×3 matrix, we can express the system $x = 3t - 4u - v$, $y = -t + v$, $z = 7t - 3u - v$ as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 & -4 & -1 \\ -1 & 0 & 1 \\ 7 & -3 & -1 \end{pmatrix}\begin{pmatrix} t \\ u \\ v \end{pmatrix}.$$

(c) We'll take two different approaches to find $\mathbf{AB}$.

Method 1. If we substitute the expression we found for $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ in part (b) into the equation we found in part (a), we have

$$\begin{pmatrix} 3 & 5 & -2 \\ 0 & 1 & -3 \\ -2 & 4 & 1 \end{pmatrix}\begin{pmatrix} 3 & -4 & -1 \\ -1 & 0 & 1 \\ 7 & -3 & -1 \end{pmatrix}\begin{pmatrix} t \\ u \\ v \end{pmatrix} = \begin{pmatrix} p \\ q \\ r \end{pmatrix}. \tag{11.1}$$

Equivalently, we can substitute the given equations $x = 3t - 4u - v$, $y = -t + v$, $z = 7t - 3u - v$ into the original system of equations to get

$$\begin{aligned} 3(3t - 4u - v) + 5(-t + v) - 2(7t - 3u - v) &= p, \\ (-t + v) - 3(7t - 3u - v) &= q, \\ -2(3t - 4u - v) + 4(-t + v) + (7t - 3u - v) &= r. \end{aligned}$$

Simplifying these equations gives

$$\begin{aligned} -10t - 6u + 4v &= p, \\ -22t + 9u + 4v &= q, \\ -3t + 5u + 5v &= r, \end{aligned}$$

which we can write as

$$\begin{pmatrix} -10 & -6 & 4 \\ -22 & 9 & 4 \\ -3 & 5 & 5 \end{pmatrix} \begin{pmatrix} t \\ u \\ v \end{pmatrix} = \begin{pmatrix} p \\ q \\ r \end{pmatrix}.$$

Comparing this to Equation (11.1), we have

$$\begin{pmatrix} 3 & 5 & -2 \\ 0 & 1 & -3 \\ -2 & 4 & 1 \end{pmatrix} \begin{pmatrix} 3 & -4 & -1 \\ -1 & 0 & 1 \\ 7 & -3 & -1 \end{pmatrix} \begin{pmatrix} t \\ u \\ v \end{pmatrix} = \begin{pmatrix} -10 & -6 & 4 \\ -22 & 9 & 4 \\ -3 & 5 & 5 \end{pmatrix} \begin{pmatrix} t \\ u \\ v \end{pmatrix}, \tag{11.2}$$

from which we conclude that

$$\begin{pmatrix} 3 & 5 & -2 \\ 0 & 1 & -3 \\ -2 & 4 & 1 \end{pmatrix} \begin{pmatrix} 3 & -4 & -1 \\ -1 & 0 & 1 \\ 7 & -3 & -1 \end{pmatrix} = \begin{pmatrix} -10 & -6 & 4 \\ -22 & 9 & 4 \\ -3 & 5 & 5 \end{pmatrix}.$$

Method 2. We use the definition of the product of a matrix and a vector to find **AB**. Since $\mathbf{B}\begin{pmatrix} t \\ u \\ v \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, we have

$$(\mathbf{AB})\begin{pmatrix} t \\ u \\ v \end{pmatrix} = \mathbf{A}\left(\mathbf{B}\begin{pmatrix} t \\ u \\ v \end{pmatrix}\right) = \mathbf{A}\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

In order to find the matrix that equals **AB**, we must express $\mathbf{A}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ as a single matrix times $\begin{pmatrix} t \\ u \\ v \end{pmatrix}$. We know how to multiply a vector by a matrix, so we express $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ in terms of vectors. From the system $x = 3t - 4u - v$, $y = -t + v$, $z = 7t - 3u - v$, we have:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 7 \end{pmatrix} t + \begin{pmatrix} -4 \\ 0 \\ -3 \end{pmatrix} u + \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} v.$$

Substituting this into $\mathbf{A}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ gives

$$\begin{aligned}
\begin{pmatrix} 3 & 5 & -2 \\ 0 & 1 & -3 \\ -2 & 4 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 3 & 5 & -2 \\ 0 & 1 & -3 \\ -2 & 4 & 1 \end{pmatrix} \left(\begin{pmatrix} 3 \\ -1 \\ 7 \end{pmatrix} t + \begin{pmatrix} -4 \\ 0 \\ -3 \end{pmatrix} u + \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} v \right) \\
&= \begin{pmatrix} 3 & 5 & -2 \\ 0 & 1 & -3 \\ -2 & 4 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \\ 7 \end{pmatrix} t + \begin{pmatrix} 3 & 5 & -2 \\ 0 & 1 & -3 \\ -2 & 4 & 1 \end{pmatrix} \begin{pmatrix} -4 \\ 0 \\ -3 \end{pmatrix} u + \begin{pmatrix} 3 & 5 & -2 \\ 0 & 1 & -3 \\ -2 & 4 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} v \\
&= \begin{pmatrix} -10 \\ -22 \\ -3 \end{pmatrix} t + \begin{pmatrix} -6 \\ 9 \\ 5 \end{pmatrix} u + \begin{pmatrix} 4 \\ 4 \\ 5 \end{pmatrix} v \\
&= \begin{pmatrix} -10t - 6u + 4v \\ -22t + 9u + 4v \\ -3t + 5u + 5v \end{pmatrix} \\
&= \begin{pmatrix} -10 & -6 & 4 \\ -22 & 9 & 4 \\ -3 & 5 & 5 \end{pmatrix} \begin{pmatrix} t \\ u \\ v \end{pmatrix}.
\end{aligned}$$

Therefore, we have

$$(\mathbf{AB})\begin{pmatrix} t \\ u \\ v \end{pmatrix} = \mathbf{A}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -10 & -6 & 4 \\ -22 & 9 & 4 \\ -3 & 5 & 5 \end{pmatrix}\begin{pmatrix} t \\ u \\ v \end{pmatrix},$$

from which we conclude that $\mathbf{AB} = \begin{pmatrix} 3 & 5 & -2 \\ 0 & 1 & -3 \\ -2 & 4 & 1 \end{pmatrix}\begin{pmatrix} 3 & -4 & -1 \\ -1 & 0 & 1 \\ 7 & -3 & -1 \end{pmatrix} = \begin{pmatrix} -10 & -6 & 4 \\ -22 & 9 & 4 \\ -3 & 5 & 5 \end{pmatrix}$.

□

We're now ready to find a method to multiply any two 3×3 matrices.

Problem 11.8: Find the product

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}.$$

Solution for Problem 11.8: We use part (c) of the previous problem as a guide. Let the matrices be **A** and **B**; we seek a matrix that equals the product **AB**, where

$$(\mathbf{AB})\begin{pmatrix} t \\ u \\ v \end{pmatrix} = \mathbf{A}\left(\mathbf{B}\begin{pmatrix} t \\ u \\ v \end{pmatrix}\right) \tag{11.3}$$

for all $\begin{pmatrix} t \\ u \\ v \end{pmatrix}$. We have

$$\mathbf{B}\begin{pmatrix} t \\ u \\ v \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}\begin{pmatrix} t \\ u \\ v \end{pmatrix} = \begin{pmatrix} b_{11}t + b_{12}u + b_{13}v \\ b_{21}t + b_{22}u + b_{23}v \\ b_{31}t + b_{32}u + b_{33}v \end{pmatrix} = \begin{pmatrix} b_{11} \\ b_{21} \\ b_{31} \end{pmatrix}t + \begin{pmatrix} b_{12} \\ b_{22} \\ b_{32} \end{pmatrix}u + \begin{pmatrix} b_{13} \\ b_{23} \\ b_{33} \end{pmatrix}v.$$

Substituting this into the right side of Equation (11.3) gives

$$\begin{aligned}
\mathbf{A}\left(\mathbf{B}\begin{pmatrix} t \\ u \\ v \end{pmatrix}\right) &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\left(\begin{pmatrix} b_{11} \\ b_{21} \\ b_{31} \end{pmatrix}t + \begin{pmatrix} b_{12} \\ b_{22} \\ b_{32} \end{pmatrix}u + \begin{pmatrix} b_{13} \\ b_{23} \\ b_{33} \end{pmatrix}v\right) \\
&= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} b_{11} \\ b_{21} \\ b_{31} \end{pmatrix}t + \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} b_{12} \\ b_{22} \\ b_{32} \end{pmatrix}u + \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} b_{13} \\ b_{23} \\ b_{33} \end{pmatrix}v \\
&= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} \end{pmatrix}t + \begin{pmatrix} a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \\ a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} \end{pmatrix}u + \begin{pmatrix} a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} \\ a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} \\ a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33} \end{pmatrix}v \\
&= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} & a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33} \end{pmatrix}\begin{pmatrix} t \\ u \\ v \end{pmatrix}.
\end{aligned}$$

Comparing this to Equation (11.3), we have the desired definition of the multiplication of two 3×3 matrices:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} & a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33} \end{pmatrix}.$$

Just as with the product of two 2×2 matrices, this definition is much easier to remember than it looks. We can view each row of a matrix as a vector, and each column of a matrix as a vector. The entry in row i and column j of the product **AB** is the dot product of the i^{th} row of **A** and the j^{th} column of **B**. For example, the entry in row 2 and column 3 of **AB** is

$$\begin{pmatrix} a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} b_{13} \\ b_{23} \\ b_{33} \end{pmatrix} = a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33}.$$

If we let the rows of **A** be the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$, and the columns of **B** be the vectors $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$, we can therefore express the product **AB** as follows:

$$\mathbf{AB} = \begin{pmatrix} — \mathbf{a}_1 — \\ — \mathbf{a}_2 — \\ — \mathbf{a}_3 — \end{pmatrix} \begin{pmatrix} | & | & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \\ | & | & | \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \mathbf{a}_1 \cdot \mathbf{b}_3 \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \mathbf{a}_2 \cdot \mathbf{b}_3 \\ \mathbf{a}_3 \cdot \mathbf{b}_1 & \mathbf{a}_3 \cdot \mathbf{b}_2 & \mathbf{a}_3 \cdot \mathbf{b}_3 \end{pmatrix}.$$

We can also remember the entry in row i and column j of **AB** as the sum $a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j}$. That is, this entry equals the sum of the three terms of the form $a_{ik}b_{kj}$, for $k = 1, 2$, and 3.

Finally, we can also remember this matrix product with the same method we used to find it. Each column of **AB** equals **A** times the corresponding column of **B**:

$$\mathbf{AB} = \mathbf{A} \begin{pmatrix} | & | & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \\ | & | & | \end{pmatrix} = \begin{pmatrix} | & | & | \\ \mathbf{Ab}_1 & \mathbf{Ab}_2 & \mathbf{Ab}_3 \\ | & | & | \end{pmatrix}.$$

□

Problem 11.9: Find the following products:

(a) $\begin{pmatrix} 2 & -3 & 0 \\ -1 & -2 & 5 \\ -2 & -5 & 1 \end{pmatrix} \begin{pmatrix} 4 & -1 & 2 \\ -1 & 1 & -2 \\ 5 & 0 & -2 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & -1 & 4 \\ 0 & 1 & 7 \\ 2 & -1 & 3 \end{pmatrix} \begin{pmatrix} -1 & 4 & 3 \\ 3 & 5 & -1 \\ 1 & 1 & 5 \end{pmatrix}$

Solution for Problem 11.9: We apply the rule for matrix multiplication we just learned to get the products below.

(a) $\begin{pmatrix} (2)(4) + (-3)(-1) + (0)(5) & (2)(-1) + (-3)(1) + (0)(0) & (2)(2) + (-3)(-2) + (0)(-2) \\ (-1)(4) + (-2)(-1) + (5)(5) & (-1)(-1) + (-2)(1) + (5)(0) & (-1)(2) + (-2)(-2) + (5)(-2) \\ (-2)(4) + (-5)(-1) + (1)(5) & (-2)(-1) + (-5)(1) + (1)(0) & (-2)(2) + (-5)(-2) + (1)(-2) \end{pmatrix} = \begin{pmatrix} 11 & -5 & 10 \\ 23 & -1 & -8 \\ 2 & -3 & 4 \end{pmatrix}.$

(b) $\begin{pmatrix} (3)(-1) + (-1)(3) + (4)(1) & (3)(4) + (-1)(5) + (4)(1) & (3)(3) + (-1)(-1) + (4)(5) \\ (0)(-1) + (1)(3) + (7)(1) & (0)(4) + (1)(5) + (7)(1) & (0)(3) + (1)(-1) + (7)(5) \\ (2)(-1) + (-1)(3) + (3)(1) & (2)(4) + (-1)(5) + (3)(1) & (2)(3) + (-1)(-1) + (3)(5) \end{pmatrix} = \begin{pmatrix} -2 & 11 & 30 \\ 10 & 12 & 34 \\ -2 & 6 & 22 \end{pmatrix}$

□

Problem 11.10: If **A** and **B** are 3×3 matrices, must we have **AB** = **BA**?

Solution for Problem 11.10: We test if **AB** and **BA** must be equal by reversing the order of the matrices in the first product we computed in the previous problem. So, we will compare the product

$$\begin{pmatrix} 4 & -1 & 2 \\ -1 & 1 & -2 \\ 5 & 0 & -2 \end{pmatrix} \begin{pmatrix} 2 & -3 & 0 \\ -1 & -2 & 5 \\ -2 & -5 & 1 \end{pmatrix}$$

to the product we found in part (a) of Problem 11.9. Applying the definition of matrix multiplication, we find that

this product equals

$$\begin{pmatrix} (4)(2)+(-1)(-1)+(2)(-2) & (4)(-3)+(-1)(-2)+(2)(-5) & (4)(0)+(-1)(5)+(2)(1) \\ (-1)(2)+(1)(-1)+(-2)(-2) & (-1)(-3)+(1)(-2)+(-2)(-5) & (-1)(0)+(1)(5)+(-2)(1) \\ (5)(2)+(0)(-1)+(-2)(-2) & (5)(-3)+(0)(-2)+(-2)(-5) & (5)(0)+(0)(5)+(-2)(1) \end{pmatrix} = \begin{pmatrix} 5 & -20 & -3 \\ 1 & 11 & 3 \\ 14 & -5 & -2 \end{pmatrix}.$$

This does not match the result we found in part (a) of Problem 11.9, so we conclude that **AB** and **BA** are not necessarily equal. □

WARNING!! Matrix multiplication is *not* commutative. That is, if **A** and **B** are matrices, then the products **AB** and **BA** are not necessarily equal.

While 3×3 matrix multiplication is not commutative, it is associative and distributive.

Important: For any 3×3 matrices **A**, **B**, and **C**, we have

$$\mathbf{(AB)C = A(BC)},$$
$$\mathbf{A(B + C) = AB + AC},$$
$$\mathbf{(A + B)C = AC + BC}.$$

Because matrix multiplication is associative, we can write a product such as **A(BC)** as simply **ABC**.

The algebraic proofs of these properties are cumbersome and not terribly illuminating, so we won't include them here.

Sidenote: Just as with 2×2 matrices, we can see that 3×3 matrix multiplication is not commutative, but is associative, by thinking of matrix multiplication as the composition of the functions represented by the matrices.

Problem 11.11: If **A** and **B** are 3×3 matrices such that **Ai = Bi**, **Aj = Bj**, and **Ak = Bk**, then must **A = B**?

Solution for Problem 11.11: We have

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}.$$

Therefore, **Ai** equals the first column of **A**. The equation **Ai = Bi** then tells us that the first column of **A** is the same as the first column of **B**. Similarly, **Aj** equals the second column of **A** and **Ak** equals the third column of **A**, so **Aj = Bj** and **Ak = Bk** together tell us that the last two columns of **A** match the last two columns of **B**. Since each column of **A** matches the corresponding column of **B**, we have **A = B**. □

Our work in Problem 11.11 gives us another way to think about matrices. The first column of a matrix **A** tells us what **Ai** is, the second column tells us what **Aj** is, and the third column tells us what **Ak** is. Let's put this observation to work.

Problem 11.12:

(a) Find a 3×3 matrix **I** such that **IA = A** for any 3×3 matrix **A**.

(b) Do we also have **AI = A** for this matrix **I**?

(c) Is the matrix you found in part (a) the only matrix **I** such that **IA = A** for all **A**? Is it the only matrix such that **AI = A** for all **A**?

Solution for Problem 11.12:

(a) The first column of **IA** equals **I** times the first column of **A**. Since **IA** = **A**, the product of **I** and the first column of **A** must therefore be the first column of **A**. That is, we must have **Iv** = **v** for any vector **v**. Now, we can find **I** by letting **v** be each of **i**, **j**, and **k**. Since **Ii** = **i**, the first column of **I** is **i**. Similarly, since **Ij** = **j** and **Ik** = **k**, the second and third columns of **I** are **j** and **k**.

So, if there is a matrix **I** such that **IA** = **A** for all **A**, then it must be the matrix whose columns are **i**, **j**, and **k**, in that order. Checking, we see that we do indeed have

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

so **IA** = **A** for any 3×3 matrix **A**.

(b) We have

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

so the matrix **I** we found in part (a) satisfies **IA** = **A** for any **A**.

(c) Our work in part (a) shows that if **IA** = **A** for all **A**, then the columns of **I** must be **i**, **j**, and **k**, in that order. So, we have shown that this matrix **I** is unique.

To see that this matrix is also the only matrix **I** such that **AI** = **A** for all **A**, suppose we have some matrix **J** such that **AJ** = **A** for all **A**. Now, we let **A** = **I**, the matrix we found in part (a). Then, we have **IJ** = **I**. But we have **IJ** = **J** because **IA** = **A** for any **A**, so we must have **J** = **I**. Therefore, the matrix **I** that we found in part (a) is the unique matrix such that **AI** = **A** for all **A**.

□

Important: The 3×3 matrix **I** such that **AI** = **IA** = **A** for any 3×3 matrix **A** is

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We call this the **identity matrix**.

Problem 11.13:

(a) Find the matrix **P** such that for every **A**, the product **PA** results in a matrix whose third row is the same as that of **A**, but the first two rows are swapped. (That is, the second row of the product is the first row of **A** and vice-versa.)

(b) Find a matrix **T** such that $\mathbf{T} \neq \mathbf{I}$ and $\mathbf{T}^3 = \mathbf{I}$.

Solution for Problem 11.13:

(a) We find **P** the same way we found the identity matrix, by finding **Pi**, **Pj**, and **Pk**. Since each column of **PA** equals **P** times the corresponding column of **A**, multiplying a vector by **P** exchanges the first and second rows of the vector but leaves the last row unchanged. Exchanging the first and second rows of **i** gives **j**, so **Pi** = **j**, which means the first column of **P** is **j**. Exchanging first and second rows of **j** gives **i**, so **Pj** = **i**, which means the second column of **P** is **i**. Finally, **Pk** = **k**, so the third column of **P** is **k**. Checking, we see

that we do indeed have

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

(b) We could write out a 3×3 matrix for $\mathbf{T}$, cube it, and then compare it to $\mathbf{I}$, but that looks pretty terrifying. Trial-and-error doesn't seem like fun either—who wants to cube random 3×3 matrices?!?

Instead, we use part (a) as inspiration. There, we found a matrix $\mathbf{P}$ such that the product $\mathbf{PA}$ results in two rows of $\mathbf{A}$ being swapped. So, consider the product $\mathbf{PPA}$. The product $\mathbf{PA}$ is just $\mathbf{A}$ with the first two rows swapped. Multiplying again by $\mathbf{P}$ to produce $\mathbf{P}^2\mathbf{A}$ simply swaps those two rows again, and gives us $\mathbf{A}$ back! So, $\mathbf{P}^2\mathbf{A} = \mathbf{A}$ for any $\mathbf{A}$, which means $\mathbf{P}^2 = \mathbf{I}$ (as we showed in Problem 11.12).

But we want $\mathbf{T}^3 = \mathbf{I}$, not $\mathbf{T}^2 = \mathbf{I}$. Rather than swapping two rows, what if we cycled them? In other words, perhaps we can find a matrix $\mathbf{T}$ such that multiplying by $\mathbf{T}$ to get $\mathbf{TA}$ moves the first two rows of $\mathbf{A}$ down one row and moves the bottom row to the top. In other words, we want the matrix $\mathbf{T}$ such that $\mathbf{Ti} = \mathbf{j}$, $\mathbf{Tj} = \mathbf{k}$, and $\mathbf{Tk} = \mathbf{i}$. These results tell us that the columns of $\mathbf{T}$ are, in order, $\mathbf{j}$, $\mathbf{k}$, and $\mathbf{i}$. Checking, we see that we have

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}.$$

Multiplying a matrix $\mathbf{A}$ by $\mathbf{T}$ moves each row of $\mathbf{A}$ down one row while the last row of $\mathbf{A}$ becomes the first row of the product $\mathbf{TA}$. Performing this operation three times gives us the original matrix back, which means that $\mathbf{T}^3\mathbf{A} = \mathbf{A}$. So, we have found a matrix $\mathbf{T}$ such that $\mathbf{T}^3$ is the identity. Let's check:

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \left(\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \right) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Success!

□

The two matrices we found in Problem 11.13 are called **permutation matrices**. A permutation matrix is a matrix $\mathbf{P}$ such that for any matrix $\mathbf{A}$, the rows of the product $\mathbf{PA}$ are just a reordering of the rows of $\mathbf{A}$ (or exactly the same ordering—the identity matrix is a permutation matrix, too).

Exercises

11.2.1 Find the following products:

(a) $\begin{pmatrix} 4 & -1 & 3 \\ -1 & 0 & 5 \\ -3 & -1 & 6 \end{pmatrix} \begin{pmatrix} -1 \\ 4 \\ -5 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & -3 & 2 \\ -1 & -2 & 7 \\ -3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -3 \\ 3 \end{pmatrix}$

11.2.2 Compute the product $\begin{pmatrix} -1 & -3 & 6 \\ -2 & 4 & 0 \\ -1 & -2 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 3 \\ 10 & -2 & -3 \\ 4 & 6 & 0.5 \end{pmatrix}$.

11.2.3 Let k be a nonzero constant. Find a matrix $\mathbf{M}$ such that $\mathbf{MA} = k\mathbf{A}$ for any matrix $\mathbf{A}$

11.2.4 Is the matrix $\mathbf{T}$ we found in part (b) of Problem 11.13 the only matrix $\mathbf{T}$ such that $\mathbf{T} \neq \mathbf{I}$ and $\mathbf{T}^3 = \mathbf{I}$?

11.3 Determinants of 3×3 Matrices

We introduced the determinant of a 2×2 matrix by solving a system of two linear equations in two variables. So, you might not be surprised that we introduce the determinant of a 3×3 matrix by solving a system of three linear equations in three variables.

Problem 11.14: Consider the system of equations

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1, \\ a_2x + b_2y + c_2z &= d_2, \\ a_3x + b_3y + c_3z &= d_3, \end{aligned}$$

where the a_i, b_i, c_i, and d_i are constants. Show that if $a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \neq 0$, then

$$x = \frac{d_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} d_2 & c_2 \\ d_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} d_2 & b_2 \\ d_3 & b_3 \end{vmatrix}}{a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}}.$$

Solution for Problem 11.14: We start by eliminating z from the second and third equations. Subtracting c_2 times the first equation from c_1 times the second equation and subtracting c_3 times the first equation from c_1 times the third gives the system

$$\begin{aligned} (a_2c_1 - a_1c_2)x + (b_2c_1 - b_1c_2)y &= d_2c_1 - d_1c_2, \\ (a_3c_1 - a_1c_3)x + (b_3c_1 - b_1c_3)y &= d_3c_1 - d_1c_3. \end{aligned}$$

We can use Cramer's rule to solve for x and y if and only if $\begin{vmatrix} a_2c_1 - a_1c_2 & b_2c_1 - b_1c_2 \\ a_3c_1 - a_1c_3 & b_3c_1 - b_1c_3 \end{vmatrix} \neq 0$. Let's take a closer look at this determinant. We have

$$\begin{vmatrix} a_2c_1 - a_1c_2 & b_2c_1 - b_1c_2 \\ a_3c_1 - a_1c_3 & b_3c_1 - b_1c_3 \end{vmatrix} = (a_2c_1 - a_1c_2)(b_3c_1 - b_1c_3) - (b_2c_1 - b_1c_2)(a_3c_1 - a_1c_3).$$

When we expand the products, the $a_1b_1c_2c_3$ terms cancel. All remaining terms have at least one factor of c_1. Factoring c_1 out of everything and grouping the remaining terms by factors a_1, b_1, and c_1, we have

$$\begin{vmatrix} a_2c_1 - a_1c_2 & b_2c_1 - b_1c_2 \\ a_3c_1 - a_1c_3 & b_3c_1 - b_1c_3 \end{vmatrix} = c_1[a_1(b_2c_3 - b_3c_2) + b_1(c_2a_3 - c_3a_2) + c_1(b_3a_2 - b_2a_3)].$$

We recognize the expressions in the inner parentheses as determinants of 2×2 matrices, and we have

$$\begin{vmatrix} a_2c_1 - a_1c_2 & b_2c_1 - b_1c_2 \\ a_3c_1 - a_1c_3 & b_3c_1 - b_1c_3 \end{vmatrix} = c_1\left(a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}\right). \tag{11.4}$$

(We'll explain why we write the middle term as $-b_1\begin{vmatrix} d_2 & c_2 \\ d_3 & c_3 \end{vmatrix}$ rather than $b_1\begin{vmatrix} c_2 & d_2 \\ c_3 & d_3 \end{vmatrix}$ later in this section.)

We need the determinant on the left side of Equation (11.4) to be nonzero in order to be able to apply Cramer's rule to the system

$$\begin{aligned} (a_2c_1 - a_1c_2)x + (b_2c_1 - b_1c_2)y &= d_2c_1 - d_1c_2, \\ (a_3c_1 - a_1c_3)x + (b_3c_1 - b_1c_3)y &= d_3c_1 - d_1c_3. \end{aligned}$$

We are given that the expression in parentheses on line (11.4) is nonzero. But we'll also need $c_1 \neq 0$ in order to apply Cramer's rule. We'll tackle the case in which $c_1 \neq 0$ now, and defer the case in which $c_1 = 0$ to Problem 11.18.

> **WARNING!!** ☢ We have to be very careful about 0, and often have to deal with 0 as a separate case when constructing proofs and solving problems.

Assuming that $c \neq 0$, the determinant on line (11.4) is nonzero, and we can apply Cramer's rule to find

$$x = \frac{\begin{vmatrix} d_2c_1 - d_1c_2 & b_2c_1 - b_1c_2 \\ d_3c_1 - d_1c_3 & b_3c_1 - b_1c_3 \end{vmatrix}}{\begin{vmatrix} a_2c_1 - a_1c_2 & b_2c_1 - b_1c_2 \\ a_3c_1 - a_1c_3 & b_3c_1 - b_1c_3 \end{vmatrix}}.$$

Following the same expansion, cancellation, and factorization steps as above, we find that

$$x = \frac{c_1[d_1(b_2c_3 - b_3c_2) + b_1(c_2d_3 - c_3d_2) + c_1(b_3d_2 - b_2d_3)]}{c_1[a_1(b_2c_3 - b_3c_2) + b_1(c_2a_3 - c_3a_2) + c_1(b_3a_2 - b_2a_3)]} = \frac{d_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} d_2 & c_2 \\ d_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} d_2 & b_2 \\ d_3 & b_3 \end{vmatrix}}{a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}}.$$

As a reminder, we'll deal with the case in which $c_1 = 0$ in Problem 11.18. □

We use the unwieldy expressions in the numerator and denominator of our result for x above to define the **determinant** of a 3×3 matrix.

Definition: We define the determinant of a 3×3 matrix as follows:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

$$= a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - a_1c_2b_3 - b_1a_2c_3 - c_1b_2a_3.$$

We can also denote the determinant of a matrix $\mathbf{A}$ as $\det(\mathbf{A})$.

Letting $\mathbf{A} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$, we can write the original system in Problem 11.14 as

$$\mathbf{A}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}.$$

In Problem 11.14, we found that if $\det(\mathbf{A}) \neq 0$, then

$$x = \frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}} = \frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\det(\mathbf{A})}.$$

This gives us an expression for x that is very much like the expression we found in Cramer's rule for systems of two linear equations with two variables. Indeed, we can go through essentially the same steps to show that if

$\det(\mathbf{A}) \neq 0$, then the system has the solution

$$x = \frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\det(\mathbf{A})}, \qquad y = \frac{\begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}}{\det(\mathbf{A})}, \qquad z = \frac{\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}}{\det(\mathbf{A})}.$$

We call this **Cramer's rule** for a three-variable system of three linear equations.

Cramer's rule is just one application of determinants of 3×3 matrices. You'll encounter applications of determinants in various areas of mathematics, physics, economics, and engineering.

Looking back at our solution to Problem 11.14, we see one particularly useful application of the determinant. Specifically, our work there, together with corresponding work solving for y and z, shows that the system

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1, \\ a_2x + b_2y + c_2z &= d_2, \\ a_3x + b_3y + c_3z &= d_3, \end{aligned}$$

has a unique solution for x if

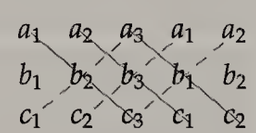
$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \neq 0.$$

In other words, if the determinant of the matrix given by the coefficients of the variables in the system is nonzero, then the system has a unique solution.

Important: For any matrix $\mathbf{A}$ and any vector $\mathbf{c}$, if $\det(\mathbf{A}) \neq 0$, then there is a unique vector $\mathbf{v}$ such that $\mathbf{Av} = \mathbf{c}$.

Just as with 2×2 matrices, we say that a 3×3 matrix is **singular** if its determinant is 0 and **nonsingular** if its determinant is nonzero. Therefore, the system $\mathbf{Av} = \mathbf{c}$ has a unique solution if $\mathbf{A}$ is nonsingular.

As you'll see in Section 12.5, we can use this observation as a starting point to determine whether or not a given 3×3 matrix has an inverse. But we'll need a few more important concepts before we can tackle that, so now we'll get more accustomed to the determinant by exploring some of its properties.



When you're first getting started with the determinants of 3×3 matrices, it can be pretty hard to remember how to compute them. One handy way of remembering the expression for the determinant of a 3×3 matrix is shown at right. We copy the first two columns of the matrix after the third column as shown. We add the products of three numbers along the three northwest-to-southeast diagonals (the diagonals along solid gray lines in the diagram) and subtract the products of three numbers along the three northeast-to-southwest diagonals (those along dashed gray lines in the diagram). This is sometimes referred to as the **Rule of Sarrus**.

Problems

Problem 11.15: Evaluate each of the following:

(a) $\begin{vmatrix} 4 & -1 & 0 \\ 5 & 7 & -2 \\ -3 & -2 & 1 \end{vmatrix}$ (b) $\begin{vmatrix} 5 & 3 & -1 \\ 4 & 1 & -5 \\ -3 & 2 & 6 \end{vmatrix}$

Problem 11.16: Let $\mathbf{A} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$. The **transpose** of a $\mathbf{A}$, which is denoted as $\mathbf{A}^T$, is defined as follows:

$$\mathbf{A}^T = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

We can think of the transpose as the result of "flipping" the matrix over the diagonal from upper left to lower right of the matrix. Show that $\det(\mathbf{A}) = \det(\mathbf{A}^T)$.

Problem 11.17:

(a) Suppose each entry in one row of $\mathbf{A}$ is multiplied by some constant k. How is the determinant of the new matrix related to $\det(\mathbf{A})$?

(b) Suppose each entry in one column of $\mathbf{A}$ is multiplied by some constant k. How is the determinant of the new matrix related to $\det(\mathbf{A})$?

(c) How is $\det(k\mathbf{A})$ related to $\det(\mathbf{A})$, where $\mathbf{A}$ is a 3×3 matrix?

Problem 11.18:

(a) Suppose $\mathbf{B}$ is the matrix formed when we swap the first and second rows of $\mathbf{A}$. How are $\det(\mathbf{A})$ and $\det(\mathbf{B})$ related?

(b) Patch the hole in our solution to Problem 11.14 by using part (a) to address the case in which $c_1 = 0$.

Problem 11.19:

(a) Show that if one row of $\mathbf{A}$ is a constant times another row, then $\det(\mathbf{A}) = 0$.

(b) If $\det(\mathbf{A}) = 0$, then must either one row of $\mathbf{A}$ be a constant times another row or one column be a constant times another column?

Problem 11.20:

(a) Show that $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = -a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}$.

(b) Show that $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} + c_3 \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$.

Problem 11.21:

(a) How are $\begin{vmatrix} 1 & 3 & 4 \\ 2 & -1 & 2 \\ 5 & 0 & 3 \end{vmatrix} + \begin{vmatrix} -3 & 2 & -1 \\ 2 & -1 & 2 \\ 5 & 0 & 3 \end{vmatrix}$ and $\begin{vmatrix} -2 & 5 & 3 \\ 2 & -1 & 2 \\ 5 & 0 & 3 \end{vmatrix}$ related?

(b) Notice that the first row of the final matrix in part (a) is the sum of the first rows of the other two matrices, and that the last two rows of all three matrices are the same. Does the relationship you found in part (a) generalize? That is, must we always have

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1' & b_1' & c_1' \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 + a_1' & b_1 + b_1' & c_1 + c_1' \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

(c) Must we have $\det(\mathbf{A}) + \det(\mathbf{B}) = \det(\mathbf{A} + \mathbf{B})$ for any matrices $\mathbf{A}$ and $\mathbf{B}$?

Problem 11.22: Let $\mathbf{A} = \begin{pmatrix} 1 & 3 & 5 \\ -1 & 4 & 2 \\ -3 & 0 & 6 \end{pmatrix}$ and let $\mathbf{B} = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 13 & 17 \\ -3 & 0 & 6 \end{pmatrix}$, where $\mathbf{B}$ is the matrix formed by adding 3 times the first row of $\mathbf{A}$ to the second row of $\mathbf{A}$ and leaving the other two rows unchanged. How are $\det(\mathbf{A})$ and $\det(\mathbf{B})$ related? Can you generalize this observation?

Problem 11.23: Evaluate each of the following:

(a) $\begin{vmatrix} 13 & -26 & 39 \\ 8 & 32 & -23 \\ 4 & 16 & -12 \end{vmatrix}$ (b) $\begin{vmatrix} 2009 & 10050 & -2010 \\ 328 & 1640 & -328 \\ -602 & -2990 & 598 \end{vmatrix}$

We start with a little practice evaluating determinants of 3×3 matrices.

Problem 11.15: Evaluate each of the following:

(a) $\begin{vmatrix} 4 & -1 & 0 \\ 5 & 7 & -2 \\ -3 & -2 & 1 \end{vmatrix}$ (b) $\begin{vmatrix} 5 & 3 & -1 \\ 4 & 1 & -5 \\ -3 & 2 & 6 \end{vmatrix}$

Solution for Problem 11.15:

(a) We have

$$\begin{vmatrix} 4 & -1 & 0 \\ 5 & 7 & -2 \\ -3 & -2 & 1 \end{vmatrix} = 4\begin{vmatrix} 7 & -2 \\ -2 & 1 \end{vmatrix} - (-1)\begin{vmatrix} 5 & -2 \\ -3 & 1 \end{vmatrix} + 0\begin{vmatrix} 5 & 7 \\ -3 & -2 \end{vmatrix}$$
$$= 4((7)(1) - (-2)(-2)) + 1((5)(1) - (-2)(-3)) + 0 = 4(3) + 1(-1) = 11.$$

(b) We have

$$\begin{vmatrix} 5 & 3 & -1 \\ 4 & 1 & -5 \\ -3 & 2 & 6 \end{vmatrix} = 5\begin{vmatrix} 1 & -5 \\ 2 & 6 \end{vmatrix} - 3\begin{vmatrix} 4 & -5 \\ -3 & 6 \end{vmatrix} + (-1)\begin{vmatrix} 4 & 1 \\ -3 & 2 \end{vmatrix}$$
$$= 5((1)(6) - (2)(-5)) - 3((4)(6) - (-5)(-3)) + (-1)((4)(2) - (1)(-3))$$
$$= 5(16) - 3(9) + (-1)(11) = 42.$$

□

Problem 11.16: Let $\mathbf{A} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$. The **transpose** of a $\mathbf{A}$, which is denoted as $\mathbf{A}^T$, is defined as follows:

$$\mathbf{A}^T = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

We can think of the transpose as the result of "flipping" the matrix over the diagonal from upper left to lower right of the matrix. Show that $\det(\mathbf{A}) = \det(\mathbf{A}^T)$.

Solution for Problem 11.16: We have

$$\det(\mathbf{A}) = a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - a_1c_2b_3 - b_1a_2c_3 - c_1b_2a_3$$

and

$$\begin{aligned}\det(\mathbf{A}^T) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} &= a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1 \\ &= a_1b_2c_3 + c_1a_2b_3 + b_1c_2a_3 - a_1c_2b_3 - b_1a_2c_3 - c_1b_2a_3 \\ &= a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - a_1c_2b_3 - b_1a_2c_3 - c_1b_2a_3,\end{aligned}$$

so $\det(\mathbf{A}) = \det(\mathbf{A}^T)$. □

Important: For any matrix $\mathbf{A}$, we have $\det(\mathbf{A}^T) = \det(\mathbf{A})$, where $\mathbf{A}^T$ is the transpose of $\mathbf{A}$.

We'll see why this fact is useful in the next problem.

Problem 11.17:

(a) Suppose each entry in one row of $\mathbf{A}$ is multiplied by some constant k. How is the determinant of the new matrix related to $\det(\mathbf{A})$?

(b) Suppose each entry in one column of $\mathbf{A}$ is multiplied by some constant k. How is the determinant of the new matrix related to $\det(\mathbf{A})$?

(c) How is $\det(k\mathbf{A})$ related to $\det(\mathbf{A})$, where $\mathbf{A}$ is a 3×3 matrix?

Solution for Problem 11.17:

(a) We have

$$\begin{aligned}\begin{vmatrix} ka_1 & kb_1 & kc_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} &= ka_1b_2c_3 + kb_1c_2a_3 + kc_1a_2b_3 - ka_1c_2b_3 - kb_1a_2c_3 - kc_1b_2a_3 \\ &= k(a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - a_1c_2b_3 - b_1a_2c_3 - c_1b_2a_3) \\ &= k\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.\end{aligned}$$

So, multiplying the first row of a matrix by k multiplies the determinant by k. Similarly, if we multiply any row or any column by k, we multiply each term in the determinant by k, thereby multiplying the whole determinant by k.

(b) Rather than going through all the computations, we can use part (a) together with what we just learned about the determinant of the transpose of a matrix. Suppose we multiply the first column of $\mathbf{A}$ by k. Since the determinant of a matrix equals the determinant of its transpose, we have

$$\det\begin{pmatrix} ka_1 & b_1 & c_1 \\ ka_2 & b_2 & c_2 \\ ka_3 & b_3 & c_3 \end{pmatrix} = \det\begin{pmatrix} ka_1 & ka_2 & ka_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} = k\det\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = k\det(\mathbf{A}^T) = k\det(\mathbf{A}),$$

where we used part (a) to "factor" k out of the first row. We can go through similar steps if any other column is multiplied by k to show that multiplying any column of a matrix multiplies its determinant by k.

(c) Multiplying a row by k multiplies the determinant by k. To produce $k\mathbf{A}$, we multiply each of the three rows by k, thereby multiplying the determinant by k three times. So, we have $\det(k\mathbf{A}) = k^3 \det(\mathbf{A})$.

□

Important: If we produce matrix **B** by multiplying a row or a column of **A** by some constant k, then we have $\det(\mathbf{B}) = k\det(\mathbf{A})$. Therefore, for any 3×3 matrix **A**, we have $\det(k\mathbf{A}) = k^3 \det(\mathbf{A})$.

Our solution to part (b) shows how useful the fact that $\det(\mathbf{A}) = \det(\mathbf{A}^T)$ can be.

Concept: Because $\det(\mathbf{A}) = \det(\mathbf{A}^T)$, for anything we prove regarding the rows and the determinant of a matrix, we can prove a corresponding fact about the columns and the determinant of a matrix.

Problem 11.18:

(a) Suppose **B** is the matrix formed when we swap the first and second rows of **A**. How are $\det(\mathbf{A})$ and $\det(\mathbf{B})$ related?

(b) Patch the hole in our solution to Problem 11.14 by using part (a) to address the case in which $c_1 = 0$.

Solution for Problem 11.18:

(a) We have

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - a_1c_2b_3 - b_1a_2c_3 - c_1b_2a_3,$$

$$\begin{vmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_2b_1c_3 + b_2c_1a_3 + c_2a_1b_3 - a_2c_1b_3 - b_2a_1c_3 - c_2b_1a_3$$

$$= b_1a_2c_3 + c_1b_2a_3 + a_1c_2b_3 - c_1a_2b_3 - a_1b_2c_3 - b_1c_2a_3.$$

We see that the terms in the second determinant are the opposites of those in the first determinant. (We have $a_1b_2c_3$ in first and $-a_1b_2c_3$ in the second, and so on.) Therefore, the second determinant is the negative of the first.

We can go through essentially the same manipulations as above to show the following:

Important: If we produce matrix **B** by swapping any two rows or any two columns of **A**, then $\det(\mathbf{B}) = -\det(\mathbf{A})$.

(b) Let $\mathbf{A} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$. In Problem 11.14 and the discussion thereafter, we showed that if $\det(\mathbf{A})$ and $c_1 \neq 0$, then the system of equations

$$a_1x + b_1y + c_1z = d_1,$$
$$a_2x + b_2y + c_2z = d_2,$$
$$a_3x + b_3y + c_3z = d_3,$$

has the unique solution

$$x = \frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\det(\mathbf{A})}, \qquad y = \frac{\begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}}{\det(\mathbf{A})}, \qquad z = \frac{\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}}{\det(\mathbf{A})}.$$

To deal with the case in which $c_1 = 0$, we note that if $c_1 = 0$, then either c_2 or c_3 must be nonzero, since otherwise we would have $\det(\mathbf{A}) = 0$ (because there is a factor of the form c_i in each term of the expansion of $\det(\mathbf{A})$).

Suppose $c_2 \neq 0$. Then, we can view our system of equations as

$$\begin{aligned} a_2x + b_2y + c_2z &= d_2, \\ a_1x + b_1y + c_1z &= d_1, \\ a_3x + b_3y + c_3z &= d_3, \end{aligned}$$

where all we have done is swapped the first and second equations. Then, our work in Problem 11.14 tells us that

$$x = \frac{\begin{vmatrix} d_2 & b_2 & c_2 \\ d_1 & b_1 & c_1 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix}}.$$

Applying the result of part (a), we swap the first and second rows of each determinant and have

$$x = \frac{\begin{vmatrix} d_2 & b_2 & c_2 \\ d_1 & b_1 & c_1 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix}} = x = \frac{-\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{-\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}} = \frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\det(\mathbf{A})},$$

as desired. We can do essentially the same for y and z, so Cramer's rule holds if $\det(\mathbf{A}) \neq 0$, $c_1 = 0$, and $c_2 \neq 0$.

Finally, we can go through essentially the same steps as above if $c_1 = c_2 = 0$ and $c_3 \neq 0$, by swapping the first and third equations in the original system of equations. We have therefore patched the hole in our solution to Problem 11.14 and shown that the system has the unique solution given above even if $c_1 = 0$.

□

Problem 11.19:

(a) Show that if one row of $\mathbf{A}$ is a constant times another row, then $\det(\mathbf{A}) = 0$.

(b) If $\det(\mathbf{A}) = 0$, then must either one row of $\mathbf{A}$ be a constant times another row or one column be a constant times another column?

Solution for Problem 11.19:

(a) If the third row is k times the second row, then we have

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ ka_2 & kb_2 & kc_2 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ kb_2 & kc_2 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ ka_2 & kc_2 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ ka_2 & kb_2 \end{vmatrix}.$$

In each of the 2×2 determinants, one row is a constant times the other, so all three of these determinants are 0, which means that the determinant of the 3×3 matrix is zero.

If the first row is a constant times the second row, then we can swap the first and third rows and use the proof above to find that the determinant of the resulting matrix is 0. The swapping of rows only reverses the sign of the determinant, so the original matrix has determinant 0 as well. Similarly, if the first row is a constant times the third row, we can swap the first and second rows and use a proof similar to that above to show that the determinant of the original matrix is 0.

Important: If one row of a 3×3 matrix is a constant times another row, or one column of the matrix is a constant times another column, then the determinant of the matrix is 0.

(b) No! Here's a counterexample:

$$\begin{vmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 1 & 1 & 3 \end{vmatrix} = (1)(1)(3) + (0)(1)(1) + (2)(0)(1) - (1)(1)(1) - (0)(0)(3) - (2)(1)(1) = 0.$$

No row is a constant times another and no column is a constant times another, and yet the determinant is still 0. We'll learn what's going on here in the next chapter.

□

Problem 11.20:

(a) Show that $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = -a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}$.

(b) Show that $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} + c_3 \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$.

Solution for Problem 11.20:

(a) The desired expression on the right side looks a lot like our definition of the determinant, except each term on the right has an entry in the second row times a 2×2 determinant, rather than an entry in the first row times a 2×2 determinant. By swapping the first two rows in the 3×3 matrix, we produce a matrix with determinant -1 times the original matrix. The new matrix has a_2, b_2, and c_2 in the first row, which gives us the desired right side when we express the determinant with our definition:

$$\begin{aligned} \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = - \begin{vmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix} &= - \left(a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} - b_2 \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} + c_2 \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \right) \\ &= -a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}, \end{aligned}$$

as desired.

(b) We wrote the expression for a 3×3 determinant in terms of 2×2 determinants by selectively factoring out first row terms when writing the expression

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1(b_2c_3 - c_2b_3) - b_1(a_2c_3 - c_2a_3) + c_1(a_2b_3 - b_2a_3)$$

If we instead expanded all these products and grouped the resulting terms by the third column entry that appears in each term, we have

$$\begin{aligned}\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} &= a_1b_2c_3 - a_1c_2b_3 - b_1a_2c_3 + b_1c_2a_3 + c_1a_2b_3 - c_1b_2a_3 \\ &= c_1(a_2b_3 - b_2a_3) - c_2(a_1b_3 - b_1a_3) + c_3(a_1b_2 - b_1a_2) \\ &= c_1\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} - c_2\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} + c_3\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.\end{aligned}$$

(We could also have solved part (a) similarly, by grouping terms with a_2, terms with b_2, and terms with c_2.)

□

Our initial definition of the determinant of a 3×3 matrix and the two equivalent expressions we found in the previous problem are examples of a process called **expansion by minors**. A **minor** of a matrix **A**, denoted as A_{ij}, is the determinant of the matrix that remains when row i and column j are omitted from matrix **A**. So, if

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

then we find A_{12} by omitting row 1 and column 2, and then taking the determinant of the remaining 2×2 matrix:

$$\begin{pmatrix} \cancel{a_{11}} & \cancel{a_{12}} & \cancel{a_{13}} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \Rightarrow \quad A_{12} = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

Similarly, we find A_{31} by omitting row 3 and column 1, and taking the determinant of the remaining 2×2 matrix:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \cancel{a_{31}} & \cancel{a_{32}} & \cancel{a_{33}} \end{pmatrix} \quad \Rightarrow \quad A_{31} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

With this notation, we can express our definition of the determinant as follows:

$$\begin{aligned}\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \\ &= a_{11}A_{11} - a_{12}A_{12} + a_{13}A_{13}. \qquad (11.5)\end{aligned}$$

Our work in part (a) above shows that we can "expand along the second row," too:

$$\begin{aligned}\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= -a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{22}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{23}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} \\ &= -a_{21}A_{21} + a_{22}A_{22} - a_{23}A_{23}. \qquad (11.6)\end{aligned}$$

By "expand along the second row," we mean that each term in our evaluation of the determinant is the product of an entry in the second row, which is a_{2j} for $j = 1$, 2, or 3, and the corresponding minor A_{2j}. As we have seen, sometimes these terms are added and sometimes they are subtracted. There's a pattern to which are added and which are subtracted: we add when the sum of the subscripts of the minor is even, and we subtract when that sum is odd.

Noting that $(-1)^{i+j}$ is -1 if $i+j$ is odd and 1 if $i+j$ is even, we can express the "expansion along the third column" in part (b) as

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} - a_{23}\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} + a_{33}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$
$$= (-1)^{1+3}a_{13}A_{13} + (-1)^{2+3}a_{23}A_{23} + (-1)^{3+3}a_{33}A_{33}. \tag{11.7}$$

That we can expand along any row as described above can be expressed as

$$\det(\mathbf{A}) = (-1)^{i+1}a_{i1}A_{i1} + (-1)^{i+2}a_{i2}A_{i2} + (-1)^{i+3}a_{i3}A_{i3}$$

for $i = 1, 2$, or 3. (The result for $i = 1$ is line (11.5) and the result for $i = 2$ is line (11.6).) Similarly, we can express the fact that we can expand along any column in this way as

$$\det(\mathbf{A}) = (-1)^{1+j}a_{1j}A_{1j} + (-1)^{2+j}a_{2j}A_{2j} + (-1)^{3+j}a_{3j}A_{3j}$$

for $j = 1, 2$, or 3. The result for $j = 3$ is line (11.7).

Sidenote: Some sources use **cofactors** of a matrix rather than minors. A cofactor of a matrix for row i and column j of a matrix $\mathbf{A}$ is $(-1)^{i+j}A_{ij}$, where A_{ij} is a minor of $\mathbf{A}$. In other words, a cofactor is formed by multiplying a minor by 1 or -1, depending on the sign of that minor in expansion by minors.

Writing out all six products in the determinant makes it clear why we can expand by minors along any row or column. We start with

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - a_1c_2b_3 - b_1a_2c_3 - c_1b_2a_3,$$

and note that each term is of the form $a_ib_jc_k$, where i, j, k are 1, 2, 3 in some order. So, each term has one factor corresponding to each column (the letters) and one factor corresponding to each row (the numbers). We can group by any number to produce an expansion along a row, such as

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_2(b_3c_1 - b_1c_3) + b_2(a_1c_3 - a_3c_1) + c_2(a_3b_1 - a_1b_3) = -a_2\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + b_2\begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} - c_2\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix},$$

or we can group by any letter to produce an expansion along a column, as we did in part (b) of Problem 11.20.

Problem 11.21:

(a) How are $\begin{vmatrix} 1 & 3 & 4 \\ 2 & -1 & 2 \\ 5 & 0 & 3 \end{vmatrix} + \begin{vmatrix} -3 & 2 & -1 \\ 2 & -1 & 2 \\ 5 & 0 & 3 \end{vmatrix}$ and $\begin{vmatrix} -2 & 5 & 3 \\ 2 & -1 & 2 \\ 5 & 0 & 3 \end{vmatrix}$ related?

(b) Notice that the first row of the final matrix in part (a) is the sum of the first rows of the other two matrices, and that the last two rows of all three matrices are the same. Does the relationship you found in part (a) generalize? That is, must we always have

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1' & b_1' & c_1' \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 + a_1' & b_1 + b_1' & c_1 + c_1' \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

(c) Must we have $\det(\mathbf{A}) + \det(\mathbf{B}) = \det(\mathbf{A} + \mathbf{B})$ for any matrices $\mathbf{A}$ and $\mathbf{B}$?

Solution for Problem 11.21:

(a) We have

$$\begin{vmatrix} 1 & 3 & 4 \\ 2 & -1 & 2 \\ 5 & 0 & 3 \end{vmatrix} = (1)(-1)(3) + (3)(2)(5) + (4)(2)(0) - (1)(2)(0) - (3)(2)(3) - (4)(-1)(5) = 29,$$

$$\begin{vmatrix} -3 & 2 & -1 \\ 2 & -1 & 2 \\ 5 & 0 & 3 \end{vmatrix} = (-3)(-1)(3) + (2)(2)(5) + (-1)(2)(0) - (-3)(2)(0) - (2)(2)(3) - (-1)(-1)(5) = 12,$$

$$\begin{vmatrix} -2 & 5 & 3 \\ 2 & -1 & 2 \\ 5 & 0 & 3 \end{vmatrix} = (-2)(-1)(3) + (5)(2)(5) + (3)(2)(0) - (-2)(2)(0) - (5)(2)(3) - (3)(-1)(5) = 41,$$

and we see that the sum of the first two equals the third.

(b) Expanding each determinant in the sum along the first row gives:

$$\begin{aligned} \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1' & b_1' & c_1' \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} &= a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \\ &\quad + a_1' \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1' \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1' \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \\ &= (a_1 + a_1') \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - (b_1 + b_1') \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + (c_1 + c_1') \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \\ &= \begin{vmatrix} a_1 + a_1' & b_1 + b_1' & c_1 + c_1' \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}. \end{aligned}$$

There's nothing special about the first row. For example, suppose two matrices have the same first and third rows. The sum of the determinants of these matrices equals the determinant of the matrix with those same first and third rows and with second row equal to the sum of the second rows of the original two matrices:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2' & b_2' & c_2' \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 + a_2' & b_2 + b_2' & c_2 + c_2' \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

There's nothing special about rows, either. If two columns in one matrix are the same as the corresponding two columns in the other, then the sum of the determinants of these matrices equals the determinant of the matrix with those same two corresponding columns and with the other column equal to the sum of the two differing columns in the original two matrices. For example, we have

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 & c_1' \\ a_2 & b_2 & c_2' \\ a_3 & b_3 & c_3' \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 + c_1' \\ a_2 & b_2 & c_2 + c_2' \\ a_3 & b_3 & c_3 + c_3' \end{vmatrix}.$$

(c) No. We can find a counterexample quickly by letting $\mathbf{A} = \mathbf{B} = \mathbf{I}$. Then, we have $\det(\mathbf{A}) = \det(\mathbf{B}) = \det(\mathbf{I}) = 1$, so $\det(\mathbf{A}) + \det(\mathbf{B}) = 2$, but

$$\det(\mathbf{A} + \mathbf{B}) = \begin{vmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{vmatrix} = 2^3 = 8.$$

So, $\det(\mathbf{A}) + \det(\mathbf{B})$ and $\det(\mathbf{A} + \mathbf{B})$ are not necessarily the same. □

WARNING!! If $\mathbf{A}$ and $\mathbf{B}$ are 3×3 matrices, then $\det(\mathbf{A}) + \det(\mathbf{B})$ and $\det(\mathbf{A} + \mathbf{B})$ are not necessarily the same.

With a great deal of perseverance, we can show that $\det(\mathbf{A})\det(\mathbf{B}) = \det(\mathbf{AB})$ for any two 3×3 matrices $\mathbf{A}$ and $\mathbf{B}$.

Important: For any two 3×3 matrices $\mathbf{A}$ and $\mathbf{B}$ we have $\det(\mathbf{A})\det(\mathbf{B}) = \det(\mathbf{AB})$.

Problem 11.22: Let $\mathbf{A} = \begin{pmatrix} 1 & 3 & 5 \\ -1 & 4 & 2 \\ -3 & 0 & 6 \end{pmatrix}$ and let $\mathbf{B} = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 13 & 17 \\ -3 & 0 & 6 \end{pmatrix}$, where $\mathbf{B}$ is the matrix formed by adding 3 times the first row of $\mathbf{A}$ to the second row of $\mathbf{A}$ and leaving the other two rows unchanged. How are $\det(\mathbf{A})$ and $\det(\mathbf{B})$ related? Can you generalize this observation?

Solution for Problem 11.22: We have

$$\det(\mathbf{A}) = (1)(4)(6) + (3)(2)(-3) + (5)(-1)(0) - (1)(2)(0) - (3)(-1)(6) - (5)(4)(-3) = 84$$

and

$$\det(\mathbf{B}) = (1)(13)(6) + (3)(17)(-3) + (5)(2)(0) - (1)(17)(0) - (3)(2)(6) - (5)(13)(-3) = 84,$$

so $\det(\mathbf{A}) = \det(\mathbf{B})$.

There doesn't appear to be anything special about $\mathbf{A}$ or $\mathbf{B}$. So we suspect that, in general, if we add a constant times one row to another row of a matrix, then the resulting matrix will have the same determinant as the original matrix.

Rather than starting from the definition of the determinant to generalize this observation, we instead combine properties of the determinant that we discovered in earlier problems. Suppose that $\mathbf{T}$ is formed by adding k times the first row of $\mathbf{M}$ to the second row of $\mathbf{M}$. From the property we discovered in Problem 11.21, we can write $\det(\mathbf{T})$ as:

$$\det(\mathbf{T}) = \begin{vmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} + km_{11} & m_{22} + km_{12} & m_{23} + km_{13} \\ m_{31} & m_{32} & m_{33} \end{vmatrix} = \begin{vmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{vmatrix} + \begin{vmatrix} m_{11} & m_{12} & m_{13} \\ km_{11} & km_{12} & km_{13} \\ m_{31} & m_{32} & m_{33} \end{vmatrix}.$$

The first determinant on the right end is just $\det(\mathbf{M})$. The second is the determinant of a matrix in which one row is a constant times another. We showed in Problem 11.19 that the determinant of such a matrix is 0, so we have $\det(\mathbf{T}) = \det(\mathbf{M})$. □

By working through essentially the same steps as we did above, we can show the following:

Important: For any matrix $\mathbf{M}$, the determinant of a matrix formed when we add a constant times any row of $\mathbf{M}$ to another row of $\mathbf{M}$ equals $\det(\mathbf{M})$. Similarly, the determinant of a matrix formed when we add a constant times any column of $\mathbf{M}$ to another column of $\mathbf{M}$ equals $\det(\mathbf{M})$.

We usually have to grind out a bunch of multiplications to evaluate the determinant of a 3×3 matrix, which is why we typically use a computer to evaluate them. But sometimes we can use the properties we have discovered in this section to evaluate a determinant pretty quickly.

Problem 11.23: Evaluate each of the following:

(a) $\begin{vmatrix} 13 & -26 & 39 \\ 8 & 32 & -23 \\ 4 & 16 & -12 \end{vmatrix}$ (b) $\begin{vmatrix} 2009 & 10050 & -2010 \\ 328 & 1640 & -328 \\ -602 & -2990 & 598 \end{vmatrix}$

Solution for Problem 11.23:

(a) Having 0s in a matrix makes finding its determinant easier since some of the products in the expression for the determinant will be 0. So, we subtract twice the third row from the second. We also can factor 13 out of the first row and 4 out of the last row, so we have:

$$\begin{vmatrix} 13 & -26 & 39 \\ 8 & 32 & -23 \\ 4 & 16 & -12 \end{vmatrix} = \begin{vmatrix} 13 & -26 & 39 \\ 0 & 0 & 1 \\ 4 & 16 & -12 \end{vmatrix} = 13 \begin{vmatrix} 1 & -2 & 3 \\ 0 & 0 & 1 \\ 4 & 16 & -12 \end{vmatrix} = 13 \cdot 4 \begin{vmatrix} 1 & -2 & 3 \\ 0 & 0 & 1 \\ 1 & 4 & -3 \end{vmatrix}$$
$$= 52((-2)(1)(1) - (1)(1)(4)) = -312.$$

(b) We factor 10 out of the second column, and add the first column to the third to give us a 0. This gives

$$\begin{vmatrix} 2009 & 10050 & -2010 \\ 328 & 1640 & -328 \\ -602 & -2990 & 598 \end{vmatrix} = 10 \begin{vmatrix} 2009 & 1005 & -2010 \\ 328 & 164 & -328 \\ -602 & -299 & 598 \end{vmatrix} = 10 \begin{vmatrix} 2009 & 1005 & -1 \\ 328 & 164 & 0 \\ -602 & -299 & -4 \end{vmatrix}.$$

Subtracting twice the middle column from the first will give us another 0:

$$10 \begin{vmatrix} 2009 & 1005 & -1 \\ 328 & 164 & 0 \\ -602 & -299 & -4 \end{vmatrix} = 10 \begin{vmatrix} -1 & 1005 & -1 \\ 0 & 164 & 0 \\ -4 & -299 & -4 \end{vmatrix}.$$

Aha! The first and last columns are the same, so the determinant equals 0.

□

Exercises

11.3.1 Compute the following two determinants:

(a) $\begin{vmatrix} 1 & 5 & 0 \\ -3 & -2 & 6 \\ -1 & 7 & -3 \end{vmatrix}$ (b) $\begin{vmatrix} -3 & 4 & 0.5 \\ 0.25 & -6 & -1 \\ 0 & 8 & 2 \end{vmatrix}$

11.3.2 Show that $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$.

11.3.3 Show that if $\mathbf{A}$ and $\mathbf{B}$ have the same second and third columns, then $\det(\mathbf{A}) + \det(\mathbf{B}) = \det(\mathbf{A} + \mathbf{B})$.

11.3.4 Evaluate $\begin{vmatrix} 30 & 51 & 20 \\ 15 & 24 & 10 \\ -46 & 28 & -29 \end{vmatrix}$ without a calculator or computer.

11.3.5 An **upper triangular matrix** is a matrix for which all the entries below the main diagonal are 0. That is, it is a matrix of the form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}.$$

What is the determinant of such a matrix?

11.4 More Than Just 2×2 and 3×3

We introduced 2×2 and 3×3 matrices with systems of equations in which the number of variables equals the number of equations. But what if the number of equations does not equal the number of variables? For example, can we use matrices and vectors to express this system:

$$w - 3x + 5y - 2z = a,$$
$$-2w + x - y + 5z = b?$$

Yes, we can:

$$\begin{pmatrix} 1 & -3 & 5 & -2 \\ -2 & 1 & -1 & 5 \end{pmatrix} \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

On the left side, we have a couple of mathematical objects we haven't seen before, a matrix with 2 rows and 4 columns, which we call a 2×4 matrix, and a vector with 4 components. The product of the 2×4 matrix and the vector in 4 dimensions gives us a vector in 2 dimensions. In other words, the 2×4 matrix represents a linear function that maps a vector in 4 dimensions to a vector in 2 dimensions.

In general, we can make a matrix with any positive integer number of rows and any positive integer number of columns, and a vector can have any positive integer number of components. In fact, the vectors we have studied so far can be viewed as a matrix with only one column. We call such a vector a **column vector**, and a matrix with only one row can be considered a **row vector**. So, for example, each of the c columns of an $r \times c$ matrix (a matrix with r rows and c columns) is itself a vector with r components, while each of the r rows is a vector with c components.

Collectively, we call the number of rows and the number of columns of a matrix the **dimensions** of a matrix. We read the dimensions $r \times c$ as "r by c".

As we have seen, we can use the product of an $r \times c$ matrix and a vector with c components to represent the expressions on one side of a system of r linear equations with c variables. In general, we define the product of an $r \times c$ matrix and a vector in c dimensions analogously to how we defined the product of a 2×2 matrix and a vector in 2 dimensions:

$$\begin{pmatrix} m_{11} & m_{12} & m_{13} & \cdots & m_{1c} \\ m_{21} & m_{22} & m_{23} & \cdots & m_{2c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{r1} & m_{r2} & m_{r3} & \cdots & m_{rc} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_c \end{pmatrix} = \begin{pmatrix} m_{11}v_1 + m_{12}v_2 + m_{13}v_3 + \cdots + m_{1c}v_c \\ m_{21}v_1 + m_{22}v_2 + m_{23}v_3 + \cdots + m_{2c}v_c \\ m_{31}v_1 + m_{32}v_2 + m_{33}v_3 + \cdots + m_{3c}v_c \\ \vdots \\ m_{r1}v_1 + m_{r2}v_2 + m_{r3}v_3 + \cdots + m_{rc}v_c \end{pmatrix}.$$

Note again that the product of an $r \times c$ matrix and vector in c dimensions is a vector in r dimensions. So, an $r \times c$ matrix represents a linear function that maps a vector in c dimensions to a vector in r dimensions.

Important: The product of an $r \times c$ matrix and a vector in n dimensions is defined if and only if $c = n$.

We define the addition of two vectors in n dimensions and scalar multiplication of vectors in n dimensions just as we did in 2 and 3 dimensions:

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix}, \qquad k\mathbf{v} = k \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} kv_1 \\ kv_2 \\ \vdots \\ kv_n \end{pmatrix}.$$

Note that we cannot add two vectors that have differing numbers of components.

The dot product and the norm of vectors in n dimensions are defined just as they are in 2 and 3 dimensions:

Important: For any vectors $\mathbf{v}$ and $\mathbf{w}$ in n dimensions, we have

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n,$$

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

Problems

Problem 11.24: In this problem, we explore matrix addition.

(a) We can define the sum of two vectors only if the two vectors have the same number of components. Show that $\mathbf{Av} + \mathbf{Bv}$ is defined if and only if $\mathbf{A}$ and $\mathbf{B}$ have the same dimensions.

(b) Let $\mathbf{A}$ and $\mathbf{B}$ be $r \times c$ matrices. We'd like to define the matrix sum $\mathbf{A} + \mathbf{B}$ such that $(\mathbf{A} + \mathbf{B})\mathbf{v} = \mathbf{Av} + \mathbf{Bv}$ for all $\mathbf{v}$. Show that such a matrix $\mathbf{A} + \mathbf{B}$ exists and explain how to compute it.

Problem 11.25: We define the matrix $k\mathbf{B}$ such that $(k\mathbf{B})\mathbf{v} = k(\mathbf{Bv})$ for all $\mathbf{v}$. Show that each entry in $k\mathbf{B}$ is k times the corresponding entry in $\mathbf{B}$.

Problem 11.26: Let $\mathbf{A}$ be an $r \times s$ matrix and $\mathbf{B}$ be a $q \times c$ matrix. Suppose there are vectors $\mathbf{v}$ and $\mathbf{w}$ such that $\mathbf{v} = \mathbf{A}(\mathbf{Bw})$.

(a) How many components must $\mathbf{w}$ have?

(b) Show that $s = q$.

(c) How many components must $\mathbf{v}$ have?

Problem 11.27: We define the product of an $r \times s$ matrix $\mathbf{A}$ and an $s \times c$ matrix $\mathbf{B}$ to be the matrix $\mathbf{AB}$ such that $(\mathbf{AB})\mathbf{v} = \mathbf{A}(\mathbf{Bv})$ for every vector $\mathbf{v}$ in c dimensions.

(a) What are the dimensions of the matrix that equals $\mathbf{AB}$?

(b) We have seen that if $\mathbf{M}$ and $\mathbf{N}$ are 3×3 matrices, then the entry in row i and column j of the product $\mathbf{MN}$ is the dot product of row i of $\mathbf{M}$ and column j of $\mathbf{N}$. Can we use the same process to find the product $\mathbf{AB}$?

Problem 11.28: Compute the following products:

(a) $\begin{pmatrix} 5 & -1 & 0 \\ -3 & 8 & -2 \end{pmatrix} \begin{pmatrix} 4 & -3 & 8 \\ 5 & -1 & 0 \\ -2 & -2 & 8 \end{pmatrix}$

(b) $\begin{pmatrix} 4 & 8 \\ -1 & 3 \\ 0 & 8 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 6 & -2 & -3 \\ -1 & -7 & 4 \end{pmatrix}$

Problem 11.24: In this problem, we explore matrix addition.

(a) We can define the sum of two vectors only if the two vectors have the same number of components. Show that $\mathbf{Av} + \mathbf{Bv}$ is defined if and only if $\mathbf{A}$ and $\mathbf{B}$ have the same dimensions.

(b) Let $\mathbf{A}$ and $\mathbf{B}$ be $r \times c$ matrices. We'd like to define the matrix sum $\mathbf{A} + \mathbf{B}$ such that $(\mathbf{A} + \mathbf{B})\mathbf{v} = \mathbf{Av} + \mathbf{Bv}$ for all $\mathbf{v}$. Show that such a matrix $\mathbf{A} + \mathbf{B}$ exists and explain how to compute it.

Solution for Problem 11.24:

(a) Suppose that $\mathbf{v}$ has c components. In order to be able to define the product $\mathbf{Av}$, the matrix $\mathbf{A}$ must have c columns. Similarly, $\mathbf{B}$ must have c columns.

Next, let $\mathbf{A}$ have r rows, so the product $\mathbf{Av}$ is a vector with r components. In order to be able to perform the vector addition $\mathbf{Av} + \mathbf{Bv}$, the vectors $\mathbf{Av}$ and $\mathbf{Bv}$ must have the same number of components. Therefore, $\mathbf{Bv}$ must have r components, which means that $\mathbf{B}$ must have r rows. Therefore, $\mathbf{B}$ must be an $r \times c$ matrix, just like $\mathbf{A}$.

(b) Since we have

$$\begin{aligned}
\mathbf{Av} + \mathbf{Bv} &= \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1c} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & a_{r3} & \cdots & a_{rc} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_c \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} & \cdots & b_{1c} \\ b_{21} & b_{22} & b_{23} & \cdots & b_{2c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{r1} & b_{r2} & b_{r3} & \cdots & b_{rc} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_c \end{pmatrix} \\
&= \begin{pmatrix} a_{11}v_1 + a_{12}v_2 + a_{13}v_3 + \cdots + a_{1c}v_c \\ a_{21}v_1 + a_{22}v_2 + a_{23}v_3 + \cdots + a_{2c}v_c \\ \vdots \\ a_{r1}v_1 + a_{r2}v_2 + a_{r3}v_3 + \cdots + a_{rc}v_c \end{pmatrix} + \begin{pmatrix} b_{11}v_1 + b_{12}v_2 + b_{13}v_3 + \cdots + b_{1c}v_c \\ b_{21}v_1 + b_{22}v_2 + b_{23}v_3 + \cdots + b_{2c}v_c \\ \vdots \\ b_{r1}v_1 + b_{r2}v_2 + b_{r3}v_3 + \cdots + b_{rc}v_c \end{pmatrix} \\
&= \begin{pmatrix} (a_{11} + b_{11})v_1 + (a_{12} + b_{12})v_2 + (a_{13} + b_{13})v_3 + \cdots + (a_{1c} + b_{1c})v_c \\ (a_{21} + b_{21})v_1 + (a_{22} + b_{22})v_2 + (a_{23} + b_{23})v_3 + \cdots + (a_{2c} + b_{2c})v_c \\ \vdots \\ (a_{r1} + b_{r1})v_1 + (a_{r2} + b_{r2})v_2 + (a_{r3} + b_{r3})v_3 + \cdots + (a_{rc} + b_{rc})v_c \end{pmatrix} \\
&= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} & \cdots & a_{1c} + b_{1c} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} & \cdots & a_{2c} + b_{2c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{r1} + b_{r1} & a_{r2} + b_{r2} & a_{r3} + b_{r3} & \cdots & a_{rc} + b_{rc} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_c \end{pmatrix},
\end{aligned}$$

we can define the matrix sum $\mathbf{A} + \mathbf{B}$ as

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} & \cdots & a_{1c} + b_{1c} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} & \cdots & a_{2c} + b_{2c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{r1} + b_{r1} & a_{r2} + b_{r2} & a_{r3} + b_{r3} & \cdots & a_{rc} + b_{rc} \end{pmatrix}.$$

□

Important: Two matrices can be added if and only if they have the same dimensions. Each entry in the sum is the sum of the corresponding entries in the two matrices.

So, addition works in the "obvious" way. How about multiplication of a matrix by a scalar?

Problem 11.25: We define the matrix equal to $k\mathbf{B}$ such that $(k\mathbf{B})\mathbf{v} = k(\mathbf{Bv})$ for all $\mathbf{v}$. Show that each entry in $k\mathbf{B}$ is k times the corresponding entry in $\mathbf{B}$.

Solution for Problem 11.25: We have

$$(k\mathbf{B})\mathbf{v} = k(\mathbf{B}\mathbf{v}) = k\left(\begin{pmatrix} b_{11} & b_{12} & b_{13} & \cdots & b_{1c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{r1} & b_{r2} & b_{r3} & \cdots & b_{rc} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_c \end{pmatrix}\right) = k\begin{pmatrix} b_{11}v_1 + b_{12}v_2 + b_{13}v_3 + \cdots + b_{1c}v_c \\ \vdots \\ b_{r1}v_1 + b_{r2}v_2 + b_{r3}v_3 + \cdots + b_{rc}v_c \end{pmatrix}$$

$$= \begin{pmatrix} kb_{11}v_1 + kb_{12}v_2 + kb_{13}v_3 + \cdots + kb_{1c}v_c \\ \vdots \\ kb_{r1}v_1 + kb_{r2}v_2 + kb_{r3}v_3 + \cdots + kb_{rc}v_c \end{pmatrix}$$

$$= \begin{pmatrix} kb_{11} & kb_{12} & kb_{13} & \cdots & kb_{1c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ kb_{r1} & kb_{r2} & kb_{r3} & \cdots & kb_{rc} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_c \end{pmatrix}$$

$$= \begin{pmatrix} kb_{11} & kb_{12} & kb_{13} & \cdots & kb_{1c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ kb_{r1} & kb_{r2} & kb_{r3} & \cdots & kb_{rc} \end{pmatrix}\mathbf{v}.$$

So, each entry in $k\mathbf{B}$ is k times the corresponding entry in $\mathbf{B}$.

Important: We multiply a matrix by a scalar by multiplying each entry in the matrix by the scalar.

□

So, addition of two $r \times c$ matrices works in the expected way, as does multiplication of any matrix by a scalar. In the following series of problems, we explore general matrix multiplication.

Problem 11.26: Let $\mathbf{A}$ be an $r \times s$ matrix and $\mathbf{B}$ be a $q \times c$ matrix. Suppose there are vectors $\mathbf{v}$ and $\mathbf{w}$ such that $\mathbf{v} = \mathbf{A}(\mathbf{B}\mathbf{w})$.

(a) How many components must $\mathbf{w}$ have?

(b) Show that $s = q$.

(c) How many components must $\mathbf{v}$ have?

Solution for Problem 11.26:

(a) The product $\mathbf{B}\mathbf{w}$ is defined if and only if the number of columns in $\mathbf{B}$ equals the number of components in $\mathbf{w}$, so $\mathbf{w}$ must have c components.

(b) Because $\mathbf{B}$ is a $q \times c$ matrix, the product $\mathbf{B}\mathbf{w}$ must be a vector in q dimensions. The product $\mathbf{A}(\mathbf{B}\mathbf{w})$ is defined if and only if the number of columns in $\mathbf{A}$ matches the number of components in the vector $\mathbf{B}\mathbf{w}$. Therefore, we must have $s = q$.

(c) As explained in part (b), the vector $\mathbf{B}\mathbf{w}$ is a vector in s dimensions. Since $\mathbf{A}$ is an $r \times s$ matrix, it maps vectors in s dimensions to vectors in r dimensions. Therefore, the product $\mathbf{A}(\mathbf{B}\mathbf{w})$ is a vector with r components.

□

As an example, suppose $\mathbf{A}$ is a 4×2 matrix and $\mathbf{B}$ is a 2×3 matrix. If vectors $\mathbf{v}$ and $\mathbf{w}$ satisfy $\mathbf{v} = \mathbf{A}(\mathbf{B}\mathbf{w})$, then

$\mathbf{v}$ has 4 components and $\mathbf{w}$ has 3 components, and the equation $\mathbf{v} = \mathbf{A}(\mathbf{B}\mathbf{w})$ looks like:

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{pmatrix} \left(\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \right).$$

Problem 11.27: We define the product of an $r \times s$ matrix $\mathbf{A}$ and an $s \times c$ matrix $\mathbf{B}$ to be the matrix $\mathbf{AB}$ such that $(\mathbf{AB})\mathbf{v} = \mathbf{A}(\mathbf{B}\mathbf{v})$ for every vector $\mathbf{v}$ in c dimensions.

(a) What are the dimensions of the matrix that equals $\mathbf{AB}$?

(b) We have seen that if $\mathbf{M}$ and $\mathbf{N}$ are 3×3 matrices, then the entry in row i and column j of the product $\mathbf{MN}$ is the dot product of row i of $\mathbf{M}$ and column j of $\mathbf{N}$. Can we use the same process to find the product $\mathbf{AB}$?

Solution for Problem 11.27:

(a) Let $\mathbf{T} = \mathbf{AB}$. The product $\mathbf{Tv}$ is defined only if the number of columns in $\mathbf{T}$ equals the number of components of $\mathbf{v}$. Therefore, $\mathbf{T}$ has c columns.

The number of components in the product $\mathbf{Tv}$ is the number of rows in $\mathbf{T}$. Since $\mathbf{Tv} = \mathbf{A}(\mathbf{Bv})$, the number of rows in $\mathbf{T}$ equals the number of components in $\mathbf{A}(\mathbf{Bv})$. The matrix $\mathbf{A}$ maps vectors in s dimensions to vectors in r dimensions, so the result of $\mathbf{A}(\mathbf{Bv})$ is a vector in r dimensions. Therefore, $\mathbf{T}$ has r rows, which means $\mathbf{T}$ is an $r \times c$ matrix.

Important: The product $\mathbf{AB}$ is defined if and only if the number of columns in $\mathbf{A}$ equals the number of rows in $\mathbf{B}$. If $\mathbf{A}$ is an $r \times s$ matrix and $\mathbf{B}$ is an $s \times c$ matrix, then the product $\mathbf{AB}$ is an $r \times c$ matrix.

(b) We discover how to multiply two matrices (when such a product is defined) in the same way we discovered the process for multiplying two 3×3 matrices. Suppose $\mathbf{A}$ is a $r \times s$ matrix and $\mathbf{B}$ is a $s \times c$ matrix and that $(\mathbf{AB})\mathbf{v} = \mathbf{A}(\mathbf{Bv})$ for all vectors $\mathbf{v}$ in c dimensions. Let the rows of $\mathbf{A}$ be $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ and the columns of $\mathbf{B}$ be $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_c$. Then, we have

$$\begin{aligned}
(\mathbf{AB})\mathbf{v} = \mathbf{A}(\mathbf{Bv}) &= \mathbf{A}\left(\begin{pmatrix} | & | & & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_c \\ | & | & & | \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_c \end{pmatrix} \right) = \mathbf{A} \begin{pmatrix} b_{11}v_1 + b_{12}v_2 + b_{13}v_3 + \cdots + b_{1c}v_c \\ b_{21}v_1 + b_{22}v_2 + b_{23}v_3 + \cdots + b_{2c}v_c \\ \vdots \\ b_{r1}v_1 + b_{r2}v_2 + b_{r3}v_3 + \cdots + b_{rc}v_c \end{pmatrix} \\
&= \mathbf{A}(v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_c\mathbf{b}_c) \\
&= v_1\mathbf{A}\mathbf{b}_1 + v_2\mathbf{A}\mathbf{b}_2 + \cdots + v_c\mathbf{A}\mathbf{b}_c \\
&= \begin{pmatrix} | & | & & | \\ \mathbf{A}\mathbf{b}_1 & \mathbf{A}\mathbf{b}_2 & \cdots & \mathbf{A}\mathbf{b}_c \\ | & | & & | \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_c \end{pmatrix} \\
&= \begin{pmatrix} | & | & & | \\ \mathbf{A}\mathbf{b}_1 & \mathbf{A}\mathbf{b}_2 & \cdots & \mathbf{A}\mathbf{b}_c \\ | & | & & | \end{pmatrix} \mathbf{v}.
\end{aligned}$$

Therefore, we have

$$\mathbf{AB} = \begin{pmatrix} | & | & & | \\ \mathbf{A}\mathbf{b}_1 & \mathbf{A}\mathbf{b}_2 & \cdots & \mathbf{A}\mathbf{b}_c \\ | & | & & | \end{pmatrix}.$$

Since for each i, we have

$$\mathbf{A}\mathbf{b}_i = \begin{pmatrix} \text{—} & \mathbf{a}_1 & \text{—} \\ \text{—} & \mathbf{a}_2 & \text{—} \\ & \vdots & \\ \text{—} & \mathbf{a}_r & \text{—} \end{pmatrix} \mathbf{b}_i = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{b}_i \\ \mathbf{a}_2 \cdot \mathbf{b}_i \\ \vdots \\ \mathbf{a}_r \cdot \mathbf{b}_i \end{pmatrix},$$

we have

$$\mathbf{AB} = \begin{pmatrix} | & | & & | \\ \mathbf{A}\mathbf{b}_1 & \mathbf{A}\mathbf{b}_2 & \cdots & \mathbf{A}\mathbf{b}_c \\ | & | & & | \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \mathbf{a}_1 \cdot \mathbf{b}_3 & \cdots & \mathbf{a}_1 \cdot \mathbf{b}_c \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \mathbf{a}_2 \cdot \mathbf{b}_3 & \cdots & \mathbf{a}_2 \cdot \mathbf{b}_c \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_r \cdot \mathbf{b}_1 & \mathbf{a}_r \cdot \mathbf{b}_2 & \mathbf{a}_r \cdot \mathbf{b}_3 & \cdots & \mathbf{a}_r \cdot \mathbf{b}_c \end{pmatrix}.$$

□

Problem 11.28: Compute the following products:

(a) $\begin{pmatrix} 5 & -1 & 0 \\ -3 & 8 & -2 \end{pmatrix} \begin{pmatrix} 4 & -3 & 8 \\ 5 & -1 & 0 \\ -2 & -2 & 8 \end{pmatrix}$ (b) $\begin{pmatrix} 4 & 8 \\ -1 & 3 \\ 0 & 8 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 6 & -2 & -3 \\ -1 & -7 & 4 \end{pmatrix}$

Solution for Problem 11.28:

(a) The product of a 2×3 matrix and a 3×3 matrix is a 2×3 matrix:

$$\begin{pmatrix} 5 & -1 & 0 \\ -3 & 8 & -2 \end{pmatrix} \begin{pmatrix} 4 & -3 & 8 \\ 5 & -1 & 0 \\ -2 & -2 & 8 \end{pmatrix}$$

$$= \begin{pmatrix} (5)(4) + (-1)(5) + (0)(-2) & (5)(-3) + (-1)(-1) + (0)(-2) & (5)(8) + (-1)(0) + (0)(8) \\ (-3)(4) + (8)(5) + (-2)(-2) & (-3)(-3) + (8)(-1) + (-2)(-2) & (-3)(8) + (8)(0) + (-2)(8) \end{pmatrix}$$

$$= \begin{pmatrix} 15 & -14 & 40 \\ 32 & 5 & -40 \end{pmatrix}$$

(b) The product of a 4×2 matrix and a 2×3 matrix is a 4×3 matrix:

$$\begin{pmatrix} 4 & 8 \\ -1 & 3 \\ 0 & 8 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 6 & -2 & -3 \\ -1 & -7 & 4 \end{pmatrix} = \begin{pmatrix} (4)(6) + (8)(-1) & (4)(-2) + (8)(-7) & (4)(-3) + (8)(4) \\ (-1)(6) + (3)(-1) & (-1)(-2) + (3)(-7) & (-1)(-3) + (3)(4) \\ (0)(6) + (8)(-1) & (0)(-2) + (8)(-7) & (0)(-3) + (8)(4) \\ (-1)(6) + (-1)(-1) & (-1)(-2) + (-1)(-7) & (-1)(-3) + (-1)(4) \end{pmatrix}$$

$$= \begin{pmatrix} 16 & -64 & 20 \\ -9 & -19 & 15 \\ -8 & -56 & 32 \\ -5 & 9 & -1 \end{pmatrix}$$

□

We say that a matrix is **square** if it has the same number of rows and columns. We can define the determinant of a square matrix with more than 3 rows by using expansion by minors in much the same manner as we did for 3×3 matrix. This is what it looks like for a 4×4 matrix:

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \\ a_4 & c_4 & d_4 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \\ a_4 & b_4 & d_4 \end{vmatrix} - d_1 \begin{vmatrix} a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{vmatrix}.$$

(Note: the "Rule of Sarrus" gimmick for remembering the formula for the determinant of a 3×3 matrix does not work for finding the determinant of larger matrices.)

If you enjoy computer programming, try writing a program for finding the determinant of a 4×4 matrix. As a bigger challenge, try writing a program for finding the determinant of any $n \times n$ matrix.

Exercises

11.4.1 Find the product $\begin{pmatrix} 3 & 4 \\ -1 & 5 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} -3 & 4 & -2 & -1 \\ 7 & 8 & -1 & 0 \end{pmatrix}$.

11.4.2 Find the product $\begin{pmatrix} 4 \\ 1 \end{pmatrix} \begin{pmatrix} -1 & 5 & -3 \end{pmatrix}$.

11.4.3 Show that if $\mathbf{v}$ and $\mathbf{w}$ are vectors in n dimensions, then $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$.

11.4.4 Show that for any scalar a and any vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in n dimensions, then we have the following:

(a) $\mathbf{v} \cdot (a\mathbf{w}) = (a\mathbf{v}) \cdot \mathbf{w} = a\mathbf{v} \cdot \mathbf{w}$

(b) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$

11.4.5 Suppose $\mathbf{A}$ is a 2×3 matrix. Is there a matrix $\mathbf{I}$ such that $\mathbf{IA} = \mathbf{A}$ for all $\mathbf{A}$? If so, what is it?

11.5 Summary

We denote a vector in three dimensions with an ordered triple $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$, which can also be written as $a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$, where $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are unit vectors from the origin in the direction of the positive x-axis, y-axis, and z-axis, respectively. Addition, scalar multiplication, and norm are defined as follows:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a+x \\ b+y \\ c+z \end{pmatrix}, \qquad k\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} ka \\ kb \\ kc \end{pmatrix}, \qquad \left\| \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right\| = \sqrt{x^2 + y^2 + z^2},$$

and the dot product is defined as

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = v_1 \cdot w_1 + v_2 \cdot w_2 + v_3 \cdot w_3.$$

The dot product has the following properties:

- $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- $\mathbf{u} \cdot (a\mathbf{v}) = a(\mathbf{u} \cdot \mathbf{v})$
- $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
- $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$
- $\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \, \|\mathbf{w}\| \cos\theta$, where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$

A 3×3 matrix represents a linear function that maps vectors in 3 dimensions to vectors in 3 dimensions. This mapping is illustrated as a product of the matrix and a vector according to the definition

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y + a_{13}z \\ a_{21}x + a_{22}y + a_{23}z \\ a_{31}x + a_{32}y + a_{33}z \end{pmatrix}.$$

For any matrix $\mathbf{A}$, any scalar c, and any vectors $\mathbf{v}$ and $\mathbf{w}$, we have $\mathbf{A}(c\mathbf{v}) = c(\mathbf{A}\mathbf{v})$ and $\mathbf{A}(\mathbf{v} + \mathbf{w}) = \mathbf{A}\mathbf{v} + \mathbf{A}\mathbf{w}$.

We add two matrices by adding the corresponding components in the matrices, and define the product of two matrices as follows:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} & a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33} \end{pmatrix}.$$

Matrix multiplication is *not* commutative, but it is associative and distributive.

Important: The 3×3 matrix $\mathbf{I}$ such that $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$ for any 3×3 matrix $\mathbf{A}$ is

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We call this the **identity matrix**.

Definition: We define the determinant of a 3×3 matrix as follows:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

$$= a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - a_1c_2b_3 - b_1a_2c_3 - c_1b_2a_3.$$

We can also denote the determinant of a matrix $\mathbf{A}$ as $\det(\mathbf{A})$.

We say a matrix is **singular** if its determinant is 0, and **nonsingular** if its determinant is nonzero. If $\det(\mathbf{A}) \neq 0$, then for any vector $\mathbf{c}$, there is unique vector $\mathbf{v}$ such that $\mathbf{Av} = \mathbf{c}$.

The determinant has the following properties:

- If we produce matrix $\mathbf{B}$ by multiplying a row or a column of $\mathbf{A}$ by some constant k, then $\det(\mathbf{B}) = k \det(\mathbf{A})$.
- For any 3×3 matrix $\mathbf{A}$ and any constant k, we have $\det(k\mathbf{A}) = k^3 \det(\mathbf{A})$.
- For any 3×3 matrix $\mathbf{A}$, swapping the rows or columns of $\mathbf{A}$ produces a matrix with determinant $(-1) \det(\mathbf{A})$.
- If one row of a 3×3 matrix is a constant times another row, or one column of the matrix is a constant times another column, then the determinant of the matrix is 0.
- For any matrix $\mathbf{M}$, the determinant of a matrix formed when we add a constant times any row of $\mathbf{M}$ to another row of $\mathbf{M}$ equals $\det(\mathbf{M})$. Similarly, the determinant of a matrix formed when we add a constant times any column of $\mathbf{M}$ to another column of $\mathbf{M}$ equals $\det(\mathbf{M})$.

A **minor** of a matrix **A**, denoted as A_{ij}, is the determinant of the matrix that remains when row i and column j are omitted from matrix **A**. We can compute the determinant of **A** as

$$\det(\mathbf{A}) = (-1)^{i+1}a_{i1}A_{i1} + (-1)^{i+2}a_{i2}A_{i2} + (-1)^{i+3}a_{i3}A_{i3}$$

for $i = 1, 2$, or 3, and as

$$\det(\mathbf{A}) = (-1)^{1+j}a_{1j}A_{1j} + (-1)^{2+j}a_{2j}A_{2j} + (-1)^{3+j}a_{3j}A_{3j}$$

for $j = 1, 2$, or 3. This technique is called **expansion by minors**.

An matrix with r rows and c columns (an $r \times c$ matrix) represents a linear function that maps vectors with c components to vectors with r components. The product of an $r \times c$ matrix and a vector with c components is defined as follows:

$$\begin{pmatrix} m_{11} & m_{12} & m_{13} & \cdots & m_{1c} \\ m_{21} & m_{22} & m_{23} & \cdots & m_{2c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{r1} & m_{r2} & m_{r3} & \cdots & m_{rc} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_c \end{pmatrix} = \begin{pmatrix} m_{11}v_1 + m_{12}v_2 + m_{13}v_3 + \cdots + m_{1c}v_c \\ m_{21}v_1 + m_{22}v_2 + m_{23}v_3 + \cdots + m_{2c}v_c \\ m_{31}v_1 + m_{32}v_2 + m_{33}v_3 + \cdots + m_{3c}v_c \\ \vdots \\ m_{r1}v_1 + m_{r2}v_2 + m_{r3}v_3 + \cdots + m_{rc}v_c \end{pmatrix}.$$

Important: The product of an $r \times c$ matrix and a vector in n dimensions is defined if and only if $c = n$.

We can define the matrix sum $\mathbf{A} + \mathbf{B}$ if and only if **A** and **B** have the same dimensions, in which case each entry in the sum $\mathbf{A} + \mathbf{B}$ is the sum of the corresponding entries in the two matrices.

Important: The product **AB** is defined if and only if the number of columns in **A** equals the number of rows in **B**. If **A** is an $r \times s$ matrix and **B** is an $s \times c$ matrix, then the product **AB** is an $r \times c$ matrix.

Let the rows of **A** be $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ and the columns of **B** be $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_c$. If the matrix product **AB** is defined, then the product is given by

$$\mathbf{AB} = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \mathbf{a}_1 \cdot \mathbf{b}_3 & \cdots & \mathbf{a}_1 \cdot \mathbf{b}_c \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \mathbf{a}_2 \cdot \mathbf{b}_3 & \cdots & \mathbf{a}_2 \cdot \mathbf{b}_c \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_r \cdot \mathbf{b}_1 & \mathbf{a}_r \cdot \mathbf{b}_2 & \mathbf{a}_r \cdot \mathbf{b}_3 & \cdots & \mathbf{a}_r \cdot \mathbf{b}_c \end{pmatrix}.$$

REVIEW PROBLEMS

Unless stated otherwise, all vectors in the Review Problems and Challenge Problems are in 3 dimensions, and all matrices are 3×3 matrices.

11.29 Let $\mathbf{u} = \begin{pmatrix} 2 \\ 3 \\ -5 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 4 \\ 5 \\ -1 \end{pmatrix}$.

(a) Find $\|\mathbf{u}\|$ and $\|\mathbf{v}\|$.

(b) Find $(2\mathbf{u} + 3\mathbf{v}) \cdot (\mathbf{u} - 3\mathbf{v})$.

11.30 Which two of the following four vectors are orthogonal: $\begin{pmatrix} 4 \\ 5 \\ -1 \end{pmatrix}, \begin{pmatrix} -3 \\ 4 \\ 2 \end{pmatrix}, \begin{pmatrix} -4 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ -3 \\ 3 \end{pmatrix}$?

11.31 Show that $(x_1y_1 + x_2y_2 + x_3y_3)^2 \leq (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2)$.

11.32

(a) Find the product $\begin{pmatrix} 4 & 5 & -2 \\ 1 & -3 & 4 \\ 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix}$.

(b) Is the matrix in part (a) the only matrix that can be multiplied by $\begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix}$ to give your answer to part (a)? If not, then find another.

(c) Is the vector $\begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix}$ from part (a) the only vector that can be multiplied by the matrix given in part (a) to get your answer to part (a)?

11.33

(a) If $\mathbf{Av}$ equals the vector $\mathbf{0}$, then must either $\mathbf{A}$ be $\mathbf{0}$ or $\mathbf{v}$ be $\mathbf{0}$?

(b) If $\mathbf{AB}$ equals the matrix $\mathbf{0}$, then must either $\mathbf{A}$ or $\mathbf{B}$ be $\mathbf{0}$?

11.34 Find the product $\begin{pmatrix} 4 & -2 & -1 \\ 0 & 5 & -1 \\ -2 & 0 & -1 \end{pmatrix} \begin{pmatrix} -3 & 2 & 6 \\ -1 & 3 & -5 \\ 7 & 0 & 3 \end{pmatrix}$.

11.35

(a) Is there a matrix $\mathbf{P}$ such that $\mathbf{PA}$ is the matrix that results when we swap the first two columns of $\mathbf{A}$ and leave the third column unchanged?

(b) Is there a matrix $\mathbf{P}$ such that $\mathbf{AP}$ is the matrix that results when we swap the first two columns of $\mathbf{A}$ and leave the third column unchanged?

11.36 If $\mathbf{Av} + \mathbf{Bv} = \mathbf{0}$ for all vectors $\mathbf{v}$, then must $\mathbf{A} = -\mathbf{B}$?

11.37 Compute the following two determinants:

(a) $\begin{vmatrix} -2 & -4 & 1 \\ -5 & 3 & 2 \\ -3 & 1 & 8 \end{vmatrix}$ (b) $\begin{vmatrix} 1/3 & 1/5 & 6 \\ 4 & -3 & 5 \\ -1 & 6 & -10 \end{vmatrix}$

11.38 Use Cramer's rule to solve the following system of equations:

$$\begin{aligned} x - 2y + 5z &= -3, \\ 2x - 5y + 3z &= 4, \\ -5x + 2y + 3z &= -1. \end{aligned}$$

11.39 Evaluate $\begin{vmatrix} x-4 & x-3 & x-2 \\ x-1 & x & x+1 \\ x+2 & x+3 & x+4 \end{vmatrix}$.

11.40

(a) As described on page 399, the matrix $\mathbf{A}^T$ is the transpose of $\mathbf{A}$. Show that if $\mathbf{A}$ and $\mathbf{B}$ are 3×3 matrices, then $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

(b) We can extend the concept of a **transpose** matrix to $r \times c$ matrices as follows. Let $\mathbf{A}$ be an $r \times c$ matrix with row vectors, in order, of $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$. The transpose of $\mathbf{A}$, denoted $\mathbf{A}^T$, is the $c \times r$ matrix whose columns, in order, are $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$. Show that if $\mathbf{AB}$ is defined, then $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

11.41 Find the product $\begin{pmatrix} 4 & 0 & 2 \\ 5 & -1 & -1 \\ 3 & -2 & 7 \\ -1 & 5 & 2 \end{pmatrix} \begin{pmatrix} 7 & -1 \\ 0 & -2 \\ -1 & 2 \end{pmatrix}$.

11.42 If $\mathbf{A}$ is 3×4 matrix, $\mathbf{C}$ is 3×6 matrix, and $\mathbf{AB} = \mathbf{C}$, then what are the dimensions of $\mathbf{B}$?

11.43

(a) Show that if the product of two square matrices is defined, then the two matrices have the same dimensions, and the product has the same dimensions as the original matrices.

(b) If the product $\mathbf{AB}$ equals a square matrix, then must $\mathbf{A}$ and $\mathbf{B}$ be square matrices?

Challenge Problems

11.44 Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors such that each vector has magnitude 1 and each vector is orthogonal to the other two. Show that if $\mathbf{r} = a\mathbf{u} + b\mathbf{v} + c\mathbf{w}$, then we must have $a = \mathbf{r}\cdot\mathbf{u}$, $b = \mathbf{r}\cdot\mathbf{v}$, and $c = \mathbf{r}\cdot\mathbf{w}$.

11.45 Let $\mathbf{u} = \mathbf{i} + \mathbf{j} + z\mathbf{k}$ and $\mathbf{v} = 2\mathbf{i} - \mathbf{j} + 3\mathbf{k}$.

(a) Find all z such that the angle between $\mathbf{u}$ and $\mathbf{v}$ is $\pi/2$.

(b) Find all z such that the angle between $\mathbf{u}$ and $\mathbf{v}$ is $\pi/3$.

(c)★ Find the maximum possible value of $\cos^2\theta$ if it is possible to find a value of z such that the angle between $\mathbf{u}$ and $\mathbf{v}$ is θ. **Hints:** 34, 33

11.46 Express $\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix}$ as a product of three polynomials in a, b, and c with nonzero degree. **Hints:** 71

11.47 We say that a matrix $\mathbf{M}$ is **skew-symmetric** if it satisfies $\mathbf{M}^T = -\mathbf{M}$. Show that if a 3×3 matrix $\mathbf{M}$ is skew-symmetric, then $\det(\mathbf{M}) = 0$.

11.48★

(a) Show that if λ is a constant such that there exists a nonzero vector $\mathbf{v}$ for which $\mathbf{Av} = \lambda\mathbf{v}$, then $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. **Hints:** 51

(b) Let $\mathbf{A} = \begin{pmatrix} -8 & 1 & 3 \\ -4 & 1 & 7 \\ 8 & 1 & -1 \end{pmatrix}$. Find three values of λ such that $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. These values are called the **eigenvalues** of $\mathbf{A}$.

(c) For each value of λ you found in part (b), find a nonzero vector $\mathbf{v}$ such that $\mathbf{Av} = \lambda\mathbf{v}$. These vectors are called **eigenvectors** of $\mathbf{A}$.

11.49 Let $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Find $\mathbf{A}^n$.

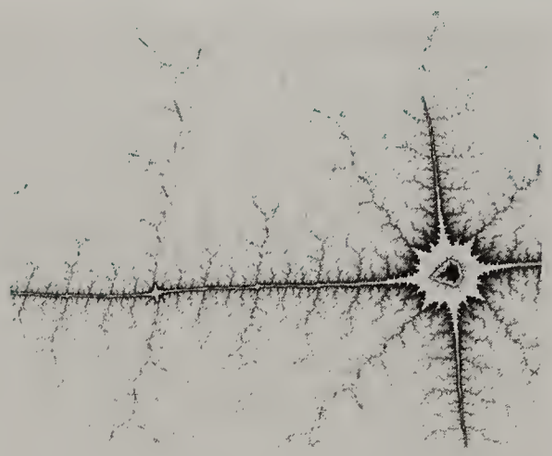
11.50 Suppose $\mathbf{A} = \begin{pmatrix} 5 & 3 & -2 \\ -1 & 5 & 4 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} -6 & -20 \\ 8 & 24 \end{pmatrix}$. Is there a matrix $\mathbf{M}$ such that $\mathbf{AM} = \mathbf{B}$? If so, is it unique?

11.51 Suppose that $\mathbf{A}$ and $\mathbf{B}$ are matrices such that $\mathbf{AB} = -\mathbf{BA}$.

(a) If $\mathbf{A}$ and $\mathbf{B}$ are 2×2 matrices, must at least one of $\mathbf{A}$ and $\mathbf{B}$ be singular?

(b) If $\mathbf{A}$ and $\mathbf{B}$ are 3×3 matrices, must at least one of $\mathbf{A}$ and $\mathbf{B}$ be singular?

11.52⋆ Let $\mathbf{A} = \begin{pmatrix} -2 & 4 & -4 \\ 0 & 2 & -6 \\ -8 & -2 & 4 \end{pmatrix}$. Prove that $\det(\mathbf{I} - \mathbf{A}^{2006})$ is not equal to 0. **Hints:** 245, 123



If there's no struggle, there's no progress. – Frederick Douglass

CHAPTER 12

Vectors and Matrices in Three Dimensions, Part 2

12.1 Lines and Planes in Three Dimensions

As in two dimensions, a **linear combination** of a set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is a sum

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$$

of products of scalars and the vectors. In this section, we'll investigate the geometry of linear combinations of vectors in three dimensions.

Problems

Problem 12.1: In this problem, we explore using equations to describe lines in three dimensions. Let A be $(3, 4, -2)$ and B be $(-1, 5, 0)$, and let ℓ be the line through A and B.

(a) Suppose that C is a point on ℓ besides A and B. How must $\vec{AC}$ and $\vec{AB}$ be related?

(b) Show that if (x, y, z) is on ℓ, then we must have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ -2 \end{pmatrix} + \begin{pmatrix} -4 \\ 1 \\ 2 \end{pmatrix} t$$

for some value of t. Does substituting any real number t into this equation produce a point (x, y, z) that is on ℓ?

Problem 12.2: Let A be $(-1, 3, 2)$, B be $(-3, 5, -1)$, C be $(0, -2, 8)$, and D be $(4, 1, 7)$. Do $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ intersect?

Problem 12.3: In this problem, we investigate linear combinations of pairs of vectors. For each of the values of $\mathbf{v}$ and $\mathbf{w}$ given below, describe the graph of all points (x, y, z) that satisfy $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$ for some values of s and t.

(a) $\mathbf{v} = \mathbf{j}$, $\mathbf{w} = 3\mathbf{j}$

(b) $\mathbf{v} = \mathbf{i}$, $\mathbf{w} = \mathbf{j}$

(c) $\mathbf{v} = \mathbf{i} + \mathbf{k}$, $\mathbf{w} = \mathbf{k}$

(d) Your answer to part (a) should be a different type of geometric object than your answers to parts (b) and (c). Why does this occur?

Problem 12.4:

(a) Describe the graph of the parametric equation $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 4 \end{pmatrix} + s\mathbf{i} + t\mathbf{j}$, where s and t are parameters.

(b) How is the graph in part (a) related to the graph of $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{i} + t\mathbf{j}$?

Problem 12.5: Suppose neither $\mathbf{v}$ nor $\mathbf{w}$ is a multiple of the other. Show that the graphs of the parametric equations

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} + s_1\mathbf{v} + t_1\mathbf{w},$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} + s_2\mathbf{v} + t_2\mathbf{w},$$

are either the same or do not intersect.

Problem 12.1: In this problem, we explore using equations to describe lines in three dimensions. Let A be $(3, 4, -2)$ and B be $(-1, 5, 0)$, and let ℓ be the line through A and B.

(a) Suppose that C is a point on ℓ besides A and B. How must $\vec{AC}$ and $\vec{AB}$ be related?

(b) Show that if (x, y, z) is on ℓ, then we must have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ -2 \end{pmatrix} + \begin{pmatrix} -4 \\ 1 \\ 2 \end{pmatrix} t$$

for some value of t. Does substituting any real number t into this equation produce a point (x, y, z) that is on ℓ?

Solution for Problem 12.1:

(a) If A, B, and C are collinear, then the angle between $\vec{AC}$ and $\vec{AB}$ must be either 0 or π, so $\vec{AC}$ must be a scalar multiple of $\vec{AB}$.

(b) Let C be (x, y, z). Since C is on ℓ, we must have $\vec{AC} = t\vec{AB}$ for some scalar t. Since $\vec{AC} = \vec{C} - \vec{A}$ and $\vec{AB} = \vec{B} - \vec{A}$, we have $\vec{C} = \vec{A} + t(\vec{B} - \vec{A})$, or

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ -2 \end{pmatrix} + \left(\begin{pmatrix} -1 \\ 5 \\ 0 \end{pmatrix} - \begin{pmatrix} 3 \\ 4 \\ -2 \end{pmatrix} \right) t = \begin{pmatrix} 3 \\ 4 \\ -2 \end{pmatrix} + \begin{pmatrix} -4 \\ 1 \\ 2 \end{pmatrix} t. \tag{12.1}$$

Any point (x, y, z) on ℓ must satisfy this equation for some value of t. Conversely, every value of t gives a point (x, y, z) on the line, because C is on ℓ whenever $\vec{AC} = t\vec{AB}$. Therefore, Equation (12.1) gives us parametric equations whose graph is a line in space: $x = 3 - 4t$, $y = 4 + t$, $z = -2 + 2t$.

□

Our solution to Problem 12.1 should be familiar, since it's the same process we used to describe a line in two dimensions in Section 9.3. We can use this same procedure to express any line in space with a parametric equation of the form

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} + t\mathbf{d}.$$

The point (a, b, c) is on the line and $\mathbf{d}$ is a **direction vector** of the line. Any vector between two distinct points on the line must be a scalar multiple of $\mathbf{d}$. We also sometimes say that this direction vector is "parallel" to the line. We can think of the parametric equation as telling us, "To get to any point (x, y, z) on the graph, start at (a, b, c) and then travel in the direction of $\mathbf{d}$ (or the opposite direction)." Since every point in the graph besides (a, b, c) is the same (or opposite) direction from (a, b, c), the graph of this equation produces the line through (a, b, c) in direction $\mathbf{d}$.

Just as in two dimensions, a given line does not have a unique direction vector, but if $\mathbf{d}_1$ and $\mathbf{d}_2$ are both possible direction vectors for the same line, then $\mathbf{d}_1 = k\mathbf{d}_2$ for some nonzero scalar k.

Problem 12.2: Let A be $(-1, 3, 2)$, B be $(-3, 5, -1)$, C be $(0, -2, 8)$, and D be $(4, 1, 7)$. Do $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ intersect?

Solution for Problem 12.2: We start by finding parametric forms for both lines. Proceeding as in the previous problem, for any point (x, y, z) on $\overleftrightarrow{AB}$, we must have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \vec{A} + (\vec{B} - \vec{A})t = \begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix} + \left(\begin{pmatrix} -3 \\ 5 \\ -1 \end{pmatrix} - \begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix} \right) t = \begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix} + \begin{pmatrix} -2 \\ 2 \\ -3 \end{pmatrix} t = \begin{pmatrix} -1 - 2t \\ 3 + 2t \\ 2 - 3t \end{pmatrix}.$$

So, $\overleftrightarrow{AB}$ is the graph of the parametric equations $x = -1 - 2t$, $y = 3 + 2t$, $z = 2 - 3t$.

Turning to $\overleftrightarrow{CD}$, the point (x, y, z) is on $\overleftrightarrow{CD}$ if

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \vec{C} + (\vec{D} - \vec{C})u = \begin{pmatrix} 0 \\ -2 \\ 8 \end{pmatrix} + \left(\begin{pmatrix} 4 \\ 1 \\ 7 \end{pmatrix} - \begin{pmatrix} 0 \\ -2 \\ 8 \end{pmatrix} \right) u = \begin{pmatrix} 0 \\ -2 \\ 8 \end{pmatrix} + \begin{pmatrix} 4 \\ 3 \\ -1 \end{pmatrix} u = \begin{pmatrix} 4u \\ -2 + 3u \\ 8 - u \end{pmatrix}.$$

Therefore, $\overleftrightarrow{CD}$ is the graph of the parametric equations $x = 4u$, $y = -2 + 3u$, $z = 8 - u$.

Since all points on $\overleftrightarrow{AB}$ are of the form $(-1 - 2t, 3 + 2t, 2 - 3t)$ for some t, and all points on $\overleftrightarrow{CD}$ are of the form $(4u, 2 + 3u, 8 - u)$ for some u, lines $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ intersect if and only if there are values of t and u for which these two points are the same. This gives us the system

$$\begin{aligned} -1 - 2t &= 4u, \\ 3 + 2t &= 2 + 3u, \\ 2 - 3t &= 8 - u. \end{aligned}$$

Adding the first equation to the second gives $2 = 2 + 7u$, from which we have $u = 0$. Substituting this into the first equation gives $t = -1/2$. However, $(t, u) = (-1/2, 0)$ does not satisfy the third equation, so there is no solution to this system of equations. Therefore, $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ do not intersect. □

Now that we have a sense of the geometry that results from considering all multiples of a single vector in three dimensions, let's take a look at linear combinations of pairs of vectors in three dimensions.

Problem 12.3: For each of the values of $\mathbf{v}$ and $\mathbf{w}$ given below, describe the graph of all points (x, y, z) that satisfy $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$ for some values of s and t.

(a) $\mathbf{v} = \mathbf{j}$, $\mathbf{w} = 3\mathbf{j}$

(b) $\mathbf{v} = \mathbf{i}$, $\mathbf{w} = \mathbf{j}$

(c) $\mathbf{v} = \mathbf{i} + \mathbf{k}$, $\mathbf{w} = \mathbf{k}$

(d) Your answer to part (a) should be a different type of geometric object than your answers to parts (b) and (c). Why does this occur?

Solution for Problem 12.3:

(a) We have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{v} + t\mathbf{w} = s\mathbf{j} + 3t\mathbf{j} = (s + 3t)\mathbf{j} = \begin{pmatrix} 0 \\ s + 3t \\ 0 \end{pmatrix}.$$

Since s and t can take on any value, the expression $s + 3t$ can take on any value. Therefore, every point on the y-axis is in the graph, and no other points are. So, the graph is the y-axis, which is a line.

(b) We have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{v} + t\mathbf{w} = s\mathbf{i} + t\mathbf{j} = \begin{pmatrix} s \\ t \\ 0 \end{pmatrix}.$$

Since s and t can take on any value, every point in the xy-plane appears in the graph. There are no points in the graph with nonzero z-coordinate, so the graph is the xy-plane.

(c) We have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{v} + t\mathbf{w} = s\mathbf{i} + s\mathbf{k} + t\mathbf{k} = s\mathbf{i} + (s + t)\mathbf{k} = \begin{pmatrix} s \\ 0 \\ s + t \end{pmatrix}.$$

To see that any point $(a, 0, b)$ in the xz-plane is in the graph, note that letting $s = a$ and $t = b - a$ in the equation above gives $(x, y, z) = (a, 0, b)$. Therefore, the graph is the xz-plane.

(d) In the first part, $\mathbf{w}$ is a multiple of $\mathbf{v}$, and the resulting graph is a line. To check if the graph is always a line when $\mathbf{w}$ is a multiple of $\mathbf{v}$, we let $\mathbf{w} = c\mathbf{v}$ for some constant c. Then, we have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{v} + t\mathbf{w} = s\mathbf{v} + tc\mathbf{v} = (s + tc)\mathbf{v}.$$

Since we can choose any values of s and t, the expression $(s + tc)\mathbf{v}$ can be any multiple of $\mathbf{v}$. Therefore, the graph is a line through the origin with direction vector $\mathbf{v}$.

In parts (b) and (c), $\mathbf{w}$ cannot be expressed as a constant times $\mathbf{v}$, and the resulting graph in each case is a plane, not a line. □

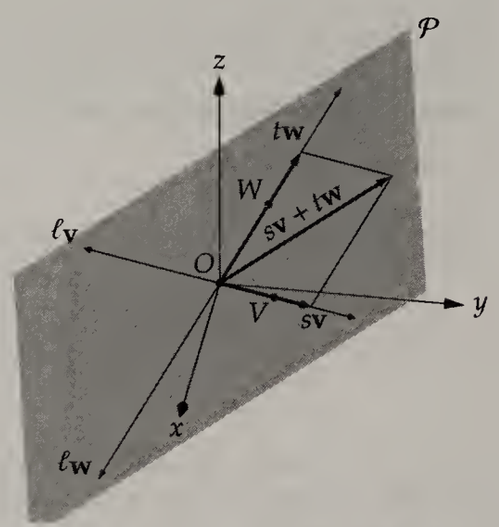
Just as we did in two dimensions, we say that a pair of vectors $\mathbf{v}$ and $\mathbf{w}$ are **linearly dependent** if and only if there are nonzero scalars a_1 and a_2 such that $a_1\mathbf{v} + a_2\mathbf{w} = \mathbf{0}$. Otherwise, the vectors are **linearly independent.** Since we can write $\mathbf{w} = c\mathbf{v}$ as $c\mathbf{v} + (-1)\mathbf{w} = \mathbf{0}$, our work in part (d) above shows that that the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{v} + t\mathbf{w}$$

is a line if $\mathbf{v}$ and $\mathbf{w}$ are linearly dependent (and not both $\mathbf{0}$).

Our work in parts (b) and (c) gives two examples in which $\mathbf{v}$ and $\mathbf{w}$ are linearly independent and the resulting graph fits our geometric concept of a plane. Let's think a little bit about why the graph of all points corresponding to linear combinations of two linearly independent vectors matches our geometric idea of a plane. What follows is not intended to be a rigorous proof, but rather to give us some intuition for why our geometric idea of a plane fits with the graph of all linear combinations of two linearly independent vectors. We'll start with the fact that any three noncollinear points in space define a plane—the plane that contains the triangle with those three points as vertices.

Suppose we start with the three noncollinear points V, W, and the origin, which we'll call O. We'll let $\mathbf{v} = \vec{V}$ and $\mathbf{w} = \vec{W}$, so $\mathbf{v}$ and $\mathbf{w}$ drawn from the origin are two sides of $\triangle VOW$. Let $\mathcal{P}$ be the plane that contains this triangle, and let $\ell_\mathbf{v}$ and $\ell_\mathbf{w}$ be the lines through the origin that contain $\mathbf{v}$ and $\mathbf{w}$, respectively. Then, lines $\ell_\mathbf{v}$ and $\ell_\mathbf{w}$ are in $\mathcal{P}$. Since any multiple of $\mathbf{v}$ lies along $\ell_\mathbf{v}$ and any multiple of $\mathbf{w}$ lies along $\ell_\mathbf{w}$, the vectors $s\mathbf{v}$ and $t\mathbf{w}$ are also in $\mathcal{P}$. We locate $s\mathbf{v} + t\mathbf{w}$ by completing the parallelogram with one vertex at the origin and sides $s\mathbf{v}$ and $t\mathbf{w}$, as shown at right. Since the triangle with sides $s\mathbf{v}$ and $t\mathbf{w}$ is in $\mathcal{P}$, so is this parallelogram, which means that $s\mathbf{v} + t\mathbf{w}$ is in $\mathcal{P}$, too.



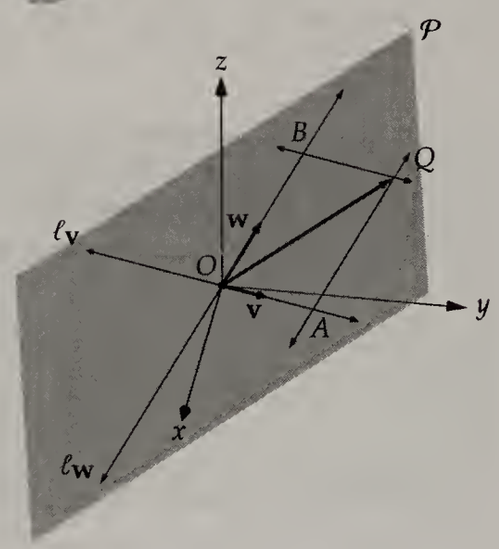
This gives us some idea why $s\mathbf{v} + t\mathbf{w}$ must be in the same plane as $\mathbf{v}$ and $\mathbf{w}$, but it doesn't explain why every point in this plane is in the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{v} + t\mathbf{w}. \tag{12.2}$$

Suppose that Q is a point in plane $\mathcal{P}$. If Q is on $\ell_\mathbf{v}$ or $\ell_\mathbf{w}$, then we clearly have $\vec{Q} = s\mathbf{v} + t\mathbf{w}$ for some s and t. Otherwise, we draw lines through Q parallel to $\ell_\mathbf{v}$ and $\ell_\mathbf{w}$, thereby completing a parallelogram $OAQB$, as shown. Since A is on $\ell_\mathbf{v}$ and B is on $\ell_\mathbf{w}$, we have $\vec{A} = s\mathbf{v}$ and $\vec{B} = t\mathbf{w}$ for some constants s and t. Therefore, we have

$$\vec{Q} = \vec{A} + \vec{B} = s\mathbf{v} + t\mathbf{w},$$

which means that Q is in the graph of (12.2).



The graph of linear combinations of two linearly independent vectors so nicely matches our concept of a plane that we use it to define a plane algebraically.

Definition: Let $\mathbf{v}$ and $\mathbf{w}$ be linearly independent. The **plane** through the origin **generated by $\mathbf{v}$ and $\mathbf{w}$** is the graph of the parametric equation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{v} + t\mathbf{w},$$

where s and t are parameters.

We will say that a vector $\mathbf{v}$ is in a plane $\mathcal{P}$ if there is some pair of points in $\mathcal{P}$ such that the vector between the points is $\mathbf{v}$. So, for example, the vector $\mathbf{i} + \mathbf{j}$ is "in the xy-plane."

Our definition only covers planes through the origin. What about planes that don't pass through the origin?

Problem 12.4:

(a) Describe the graph of the parametric equation $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 4 \end{pmatrix} + s\mathbf{i} + t\mathbf{j}$.

(b) How is the graph in part (a) related to the graph of $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{i} + t\mathbf{j}$?

Solution for Problem 12.4:

(a) We have

$$\begin{aligned} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} -1 \\ 3 \\ 4 \end{pmatrix} + s\mathbf{i} + t\mathbf{j} \\ &= \begin{pmatrix} -1 \\ 3 \\ 4 \end{pmatrix} + \begin{pmatrix} s \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} 0 \\ t \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} -1+s \\ 3+t \\ 4 \end{pmatrix}. \end{aligned}$$

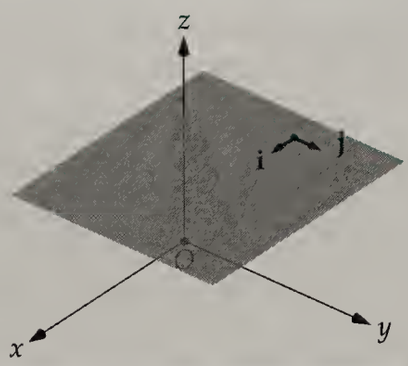
Since s and t can take on any values, the first two coordinates of $(-1+s, 3+t, 4)$ can be anything. So, the graph consists of all the points in space with $z = 4$, which means that the graph is a plane parallel to the xy-plane and 4 units above the xy-plane.

Another way to think about this is to start with the plane generated by $\mathbf{i}$ and $\mathbf{j}$. Then, we form the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 4 \end{pmatrix} + s\mathbf{i} + t\mathbf{j}$$

by adding $\begin{pmatrix} -1 \\ 3 \\ 4 \end{pmatrix}$ to every vector from the origin to a point in a plane generated by $\mathbf{i}$ and $\mathbf{j}$. In other words, we translate the plane generated by $\mathbf{i}$ and $\mathbf{j}$ by the vector $\begin{pmatrix} -1 \\ 3 \\ 4 \end{pmatrix}$.

Yet another way to think about it is that we form the graph by starting from $(-1,3,4)$, and then traveling by the vector $s\mathbf{i} + t\mathbf{j}$ to find (x,y,z). Starting from a point and then considering the result of traveling by every possible linear combination of two linearly independent vectors is exactly how we (algebraically) defined a plane through the origin. The only difference here is that we are starting from $(-1,3,4)$ instead of the origin. This observation inspires a definition of a plane defined by two linearly independent vectors and a point through which the plane passes:



Definition: Let $\mathbf{v}$ and $\mathbf{w}$ be linearly independent. The **plane** through the point (a,b,c) **generated by $\mathbf{v}$ and $\mathbf{w}$** is the graph of the parametric equation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} + s\mathbf{v} + t\mathbf{w},$$

where s and t are parameters.

Notice that when we have $(a,b,c) = (0,0,0)$, this definition matches our earlier definition of a plane through the origin.

(b) The graph of $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{i} + t\mathbf{j}$ is the xy-plane, as described in part (b) of Problem 12.3. This plane is the graph of $z = 0$. The plane in part (a) is the graph of $z = 4$. We cannot have both $z = 0$ and $z = 4$, so the plane in this part and the plane in part (a) do not intersect.

□

We say that two planes are **parallel** if either they don't intersect or they are the same plane. The planes in Problem 12.4 are generated by the same two vectors, and we found that the resulting planes are parallel. Geometrically, this makes sense. The vectors that generate a plane determine the "direction" of the plane, so it's not surprising that the two planes generated by the same vectors are parallel. We can use our algebraic definition of a plane to prove that this is the case.

Problem 12.5: Suppose $\mathbf{v}$ and $\mathbf{w}$ are linearly independent. Show that the graphs of the parametric equations

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} + s\mathbf{v} + t\mathbf{w}, \tag{12.3}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} + p\mathbf{v} + q\mathbf{w}, \tag{12.4}$$

are parallel.

Solution for Problem 12.5: The two graphs have a point in common if and only if there are scalars s_1, t_1, p_1, q_1 such that

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} + s_1\mathbf{v} + t_1\mathbf{w} = \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} + p_1\mathbf{v} + q_1\mathbf{w}. \tag{12.5}$$

If there are no such scalars, then the graphs do not intersect, and the graphs are parallel. All that's left to show is that if such scalars exist, then the graphs are the same.

For any point (x_0, y_0, z_0) on (12.3), there are scalars s_0 and t_0 such that

$$\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} + s_0\mathbf{v} + t_0\mathbf{w}. \tag{12.6}$$

To see that (x_0, y_0, z_0) is on (12.4), we first rearrange (12.5) as

$$\begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} + (s_1 - p_1)\mathbf{v} + (t_1 - q_1)\mathbf{w}. \tag{12.7}$$

Now, letting $p = p_1 - s_1 + s_0$ and $q = q_1 - t_1 + t_0$ in (12.4) gives

$$\begin{aligned} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} + (p_1 - s_1 + s_0)\mathbf{v} + (q_1 - t_1 + t_0)\mathbf{w} \\ &= \left(\begin{pmatrix} a \\ b \\ c \end{pmatrix} + (s_1 - p_1)\mathbf{v} + (t_1 - q_1)\mathbf{w} \right) + (p_1 - s_1 + s_0)\mathbf{v} + (q_1 - t_1 + t_0)\mathbf{w} \\ &= \begin{pmatrix} a \\ b \\ c \end{pmatrix} + s_0\mathbf{v} + t_0\mathbf{w}, \end{aligned}$$

which equals our expression for $\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$ from (12.6). Therefore, (x_0, y_0, z_0) is in both graphs. Similarly, we can show that any point in the graph of (12.4) is in the graph of (12.3), so the graphs of the two equations are the same.

Since the planes that are the graphs of (12.3) and (12.4) either don't intersect or are the same, they are parallel. □

Exercises

12.1.1 Find vectors $\mathbf{v}$ and $\mathbf{w}$ such that

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{v} + \mathbf{w}t$$

describes the line through $(-4, 5, 0)$ and $(-5, -1, 2)$. Are the vectors you found the only possible values of $\mathbf{v}$ and $\mathbf{w}$?

12.1.2 Find a vector that is parallel to the graph of $x = 4 - 2t$, $y = 2 + t$, $z = 2 - 3t$.

12.1.3 Let $\mathcal{P}$ be the plane through $(-2, -1, 4)$ generated by $2\mathbf{i} - \mathbf{j} + \mathbf{k}$ and $-3\mathbf{i} + 2\mathbf{k}$. Show that $(10, -4, 3)$ is on $\mathcal{P}$, but $(-1, 5, 2)$ is not.

12.1.4 Find x if $x\mathbf{i} + 7\mathbf{j} + 10\mathbf{k}$ is in the plane through the points $(-1, 2, 4)$, $(3, -1, 2)$, and $(5, 1, 6)$.

12.2 More Planes in Three Dimensions

Now that we have an algebraic definition of a plane, we can use algebraic tools on problems involving planes. In this section, we'll explore these tools. We'll also address the question of what happens when we consider linear

combinations of *three* vectors in three dimensions.

Problems

Problem 12.6: Let $\mathbf{v}$ and $\mathbf{w}$ be linearly independent and let $\mathbf{a}$ and $\mathbf{b}$ be linearly independent. In this problem, we show that if $\mathbf{v}$ and $\mathbf{w}$ are in the plane through the origin generated by $\mathbf{a}$ and $\mathbf{b}$, then the graphs of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{v} + t\mathbf{w} \qquad \text{and} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = p\mathbf{a} + q\mathbf{b}$$

are the same.

(a) Show that for every choice of s and t, we can find values of p and q such that both equations produce the same point (x, y, z). **Hints:** 32, 274

(b) Show that for every choice of p and q, we can find values of s and t such that both equations produce the same point (x, y, z). **Hints:** 5

Problem 12.7: In this problem, we find an equation whose graph is the plane through the points $(0, 4, 0)$, $(3, -1, -3)$, and $(2, -2, -4)$.

(a) Find a parametric equation whose graph is the plane.

(b) Show that there is an equation of the form $ax + by + cz = d$, where a, b, c, and d are constants, whose graph contains every point in the plane.

Problem 12.8: In this problem, we show that the graph of $n_1x + n_2y + n_3z = d$ is a plane, where n_1, n_2, n_3, and d are constants and at least one of n_1, n_2, n_3 is nonzero.

(a) Suppose $n_1 \neq 0$. Solve the equation for x. Use the result to express the equation in a form that fits our algebraic definition of a plane.

(b) Complete the proof by addressing the case $n_1 = 0$.

Problem 12.9: Let $\mathbf{v}$ and $\mathbf{w}$ be linearly independent and let x_0, y_0, and z_0 be constants.

(a) Show that we can always eliminate the parameters from the parametric equation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + s\mathbf{v} + t\mathbf{w}$$

to produce an equation of the form $n_1x + n_2y + n_3z = d$, where n_1, n_2, n_3, and d are constants.

(b) Show that every point in the graph of the equation of the form $n_1x + n_2y + n_3z = d$ produced by eliminating the parameters is in the graph of the original parametric equation.

Problem 12.10: In this problem, we find a vector that is orthogonal to every vector in the plane through the origin generated by $\mathbf{v} = 3\mathbf{i} + 4\mathbf{j} - \mathbf{k}$ and $\mathbf{w} = 4\mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$.

(a) Find an equation of the form $n_1x + n_2y + n_3z = 0$ whose graph is the plane through the origin generated by $\mathbf{v}$ and $\mathbf{w}$.

(b) Let $\mathbf{n} = n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}$. Show that $\mathbf{n}$ is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$.

(c) Is $\mathbf{n}$ orthogonal to every vector in the plane?

Problem 12.11: In Problem 12.7, you found the equation of a plane through the points $(0,4,0)$, $(3,-1,-3)$, and $(2,-2,-4)$. This answer should be of the form $n_1x + n_2y + n_3z = d$, where n_1, n_2, n_3, and d are constants. Is $\mathbf{n} = n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}$ orthogonal to the vector between any two points in the plane?

Problem 12.12: Show that the vector

$$\mathbf{n} = n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}$$

is orthogonal to the vector between any two points in the graph of

$$n_1x + n_2y + n_3z = d,$$

where n_1, n_2, n_3, and d are constants.

Problem 12.13:

(a) Find an expression for $\text{proj}_{\mathbf{w}}\mathbf{v}$ in terms of $\mathbf{v} \cdot \mathbf{w}$, $\|\mathbf{w}\|$, and $\mathbf{w}$.

(b) The distance between a point and a plane is the length of the shortest segment between the point and a point in the plane. Find an expression for the distance between the point (x_0, y_0, z_0) and the plane $ax + by + cz + d = 0$, where a, b, c, and d are constants.

Problem 12.14: Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in three dimensions such that $\mathbf{v}$ and $\mathbf{w}$ are linearly independent. We will explore the graph of the parametric equation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = r\mathbf{u} + s\mathbf{v} + t\mathbf{w}. \tag{12.8}$$

(a) Show that if $\mathbf{u}$ is a linear combination of $\mathbf{v}$ and $\mathbf{w}$, then the graph of Equation (12.8) is a plane.

(b) Suppose that $\mathbf{u}$ is not a linear combination of $\mathbf{v}$ and $\mathbf{w}$. Then, what is the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{u} + s\mathbf{v} + t\mathbf{w}?$$

(Notice that the coefficient of $\mathbf{u}$ here is 1, not r.)

(c) Suppose that $\mathbf{u}$ is not a linear combination of $\mathbf{v}$ and $\mathbf{w}$. Then, what is the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2\mathbf{u} + s\mathbf{v} + t\mathbf{w}?$$

How is your answer related to the graph in part (b)?

(d) Suppose that $\mathbf{u}$ is not a linear combination of $\mathbf{v}$ and $\mathbf{w}$. What is the graph of Equation (12.8)?

Problem 12.15: We say that $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are **linearly dependent** if there exist constants a_1, a_2, and a_3, not all 0, such that $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 = \mathbf{0}$. Otherwise, we say the vectors are **linearly independent**. Show that if $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are linearly independent, then for any (x, y, z), then there is exactly one ordered triple (r, s, t) such that

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = r\mathbf{u} + s\mathbf{v} + t\mathbf{w}.$$

Problem 12.6: Let $\mathbf{v}$ and $\mathbf{w}$ be linearly independent and let $\mathbf{a}$ and $\mathbf{b}$ be linearly independent. Show that if $\mathbf{v}$ and $\mathbf{w}$ are in the plane through the origin generated by $\mathbf{a}$ and $\mathbf{b}$, then the graphs of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{v} + t\mathbf{w} \qquad \text{and} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = p\mathbf{a} + q\mathbf{b}$$

are the same.

Solution for Problem 12.6: We must show that every point in the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{v} + t\mathbf{w}, \tag{12.9}$$

is also in the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = p\mathbf{a} + q\mathbf{b}, \tag{12.10}$$

and vice versa.

We'll first show that every point in the graph of (12.9) is on the graph of (12.10). We do so by showing that for any choice of s and t in (12.9), we can find values of p and q in (12.10) that produce the same (x, y, z).

We are given that $\mathbf{v}$ and $\mathbf{w}$ are in the graph of (12.10), so there are scalars p_1, q_1, p_2, q_2 such that

$$\mathbf{v} = p_1\mathbf{a} + q_1\mathbf{b},$$
$$\mathbf{w} = p_2\mathbf{a} + q_2\mathbf{b}.$$

Substituting these into (12.9) gives

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s(p_1\mathbf{a} + q_1\mathbf{b}) + t(p_2\mathbf{a} + q_2\mathbf{b}) = (sp_1 + tp_2)\mathbf{a} + (sq_1 + tq_2)\mathbf{b}. \tag{12.11}$$

So, for any choice of s and t in (12.9), letting $p = sp_1 + tp_2$ and $q = sq_1 + tq_2$ in (12.10) produces the same (x, y, z). Therefore, every point in the graph of (12.9) is also in the graph of (12.10).

We must now show that we can go the other way. Specifically, we must show that for any choice of p and q in (12.10), we can find s and t in (12.9) that produces the same (x, y, z). Equation (12.11) gives us a good starting point. The graph of Equation (12.11) is the same as the graph of Equation (12.9). So, if for any p and q, we can find s and t such that

$$p = sp_1 + tp_2,$$
$$q = sq_1 + tq_2,$$

then the graphs of (12.9) and (12.10) are the same. This system has a solution (s, t) for any (p, q) if and only if $\begin{vmatrix} p_1 & p_2 \\ q_1 & q_2 \end{vmatrix} \neq 0$. To show that this determinant is nonzero, we go back to the definitions of p_1, q_1, p_2, q_2. These are the scalars such that

$$\mathbf{v} = p_1\mathbf{a} + q_1\mathbf{b},$$
$$\mathbf{w} = p_2\mathbf{a} + q_2\mathbf{b}.$$

If $\begin{vmatrix} p_1 & p_2 \\ q_1 & q_2 \end{vmatrix} = 0$, then one column of $\begin{pmatrix} p_1 & p_2 \\ q_1 & q_2 \end{pmatrix}$ must be a multiple of the other. But this would mean that one of **v** and **w** is a multiple of the other, which we can't have because **v** and **w** are linearly independent. Therefore, we conclude that $\begin{vmatrix} p_1 & p_2 \\ q_1 & q_2 \end{vmatrix} \neq 0$, which means that for every pair (p, q), we can find constants s and t such that $sp_1 + tp_2 = p$ and $sq_1 + tq_2 = q$. So, for every choice of p and q in (12.10), we can find values of s and t in (12.9) that produce the same (x, y, z).

We have therefore shown that every point in the graph of (12.9) is in the graph of (12.10), and vice versa. We conclude that the graphs of these equations are the same. In other words, the plane generated by **v** and **w** is the same as the plane generated by **a** and **b**. □

As a key step in our solution to Problem 12.6, we proved that if linearly independent vectors **v** and **w** are both in the plane generated by linearly independent vectors **a** and **b**, then **a** and **b** are in the plane generated by **v** and **w**. This statement is equivalent to the following:

Important: If linearly independent vectors **v** and **w** are both linear combinations of linearly independent vectors **a** and **b**, then **a** and **b** are both linear combinations of **v** and **w**.

Since we can choose **v** and **w** to be any two linearly independent vectors in the plane generated by **a** and **b**, our work in Problem 12.6 also tells us that:

Important: Any two linearly independent vectors in a plane can be used to generate the plane.

Also, there's nothing special about planes through the origin. We could have used the same steps we used in Problem 12.6 to show that any two linearly independent vectors in a plane $\mathcal{P}$ through a point (a, b, c) generate $\mathcal{P}$.

Problem 12.7: In this problem, we find an equation whose graph is the plane through the points $(0, 4, 0)$, $(3, -1, -3)$, and $(2, -2, -4)$.

(a) Find a parametric equation whose graph is the plane.

(b) Show that there is an equation of the form $ax + by + cz = d$, where a, b, c, and d are constants, whose graph contains every point in the plane.

Solution for Problem 12.7:

(a) We know how to find equations for a plane given a point in the plane and two vectors contained by the plane. We can use the three given points to produce the two needed vectors. Let $(0, 4, 0)$ be A, let point $(3, -1, -3)$ be B, and let $(2, -2, -4)$ be C. Then, the plane is the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \vec{A} + s\vec{AB} + t\vec{AC} = \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix} + s\begin{pmatrix} 3-0 \\ -1-4 \\ -3-0 \end{pmatrix} + t\begin{pmatrix} 2-0 \\ -2-4 \\ -4-0 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix} + s\begin{pmatrix} 3 \\ -5 \\ -3 \end{pmatrix} + t\begin{pmatrix} 2 \\ -6 \\ -4 \end{pmatrix}.$$

We could have replaced $\begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}$ with $\begin{pmatrix} 3 \\ -1 \\ -3 \end{pmatrix}$, $\begin{pmatrix} 2 \\ -2 \\ -4 \end{pmatrix}$, or any other vector corresponding to a point in the plane.

(b) We can write our parametric equation as a system of equations:

$$\begin{aligned} x &= 3s + 2t, \\ y &= 4 - 5s - 6t, \\ z &= -3s - 4t. \end{aligned}$$

We produce an equation involving only x, y, and z by eliminating the parameters. Adding the first and third equations gives $x + z = -2t$. Subtracting 5 times the third equation from 3 times the second gives $3y - 5z = 12 + 2t$. Adding $x + z = -2t$ to $3y - 5z = 12 + 2t$ gives

$$x + 3y - 4z = 12.$$

□

Our solution to Problem 12.7 doesn't show that the graph of the parametric equation is the same as the graph of $x + 3y - 4z = 12$. We only showed that every point in the graph of the parametric equation is in the graph of $x + 3y - 4z = 12$. We didn't show that every point in the graph of $x + 3y - 4z = 12$ is in the graph of the parametric equations. We'll defer that proof to Problem 12.9, where we tackle the general case.

Back in Chapter 9, we showed that the graph in two dimensions of any equation of the form $ax + by = c$ is a line, and that every line in two dimensions is the graph of an equation of this form. Our work in Problem 12.7 suggests we might be able to prove a similar statement about planes in three dimensions.

In the next two problems, we show that every plane in three dimensions is the graph of an equation of the form $n_1x + n_2y + n_3z = d$. We do so by first showing that the graph of every equation of the form $n_1x + n_2y + n_3z = d$ is the same as the graph of some equation of the form

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + s\mathbf{v} + t\mathbf{w}, \tag{12.12}$$

where $\mathbf{v}$ and $\mathbf{w}$ are linearly independent. Then, we use this fact to show that the graph of any equation of the form (12.12) is the same as the graph of some equation of the form $n_1x + n_2y + n_3z = d$. Make sure you see why both of these steps are necessary—the first tells us that the graph of $n_1x + n_2y + n_3z = d$ is always a plane, and the second tells us that every plane is the graph of some equation of this form.

We go through this trouble for several reasons. Sometimes equations of the form $n_1x + n_2y + n_3z = d$ are much more convenient to work with than equations of the form (12.12). More importantly, we will use these forms together to explore graphs of linear combinations of three linearly independent vectors. Finally, in Section 12.3, we'll use them to discover a very useful operation for vectors in three dimensions.

Problem 12.8: Show that the graph of $n_1x + n_2y + n_3z = d$ is a plane, where n_1, n_2, n_3, and d are constants and at least one of n_1, n_2, n_3 is nonzero.

Solution for Problem 12.8: To show that the graph must be a plane, we must show that the graph fits our algebraic definition of a plane. This means we have to show that it is the graph of a parametric equation of the form

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} + s\mathbf{v} + t\mathbf{w}$$

for some linearly independent vectors $\mathbf{v}$ and $\mathbf{w}$ and some constants a, b, and c. To see how, let's focus on x in the parametric form. We wish to have $x = a + sv_1 + tw_1$, where v_1 and w_1 are the first components of $\mathbf{v}$ and $\mathbf{w}$. On the right side of the parametric equation, only s and t can vary—a, v_1, and w_1 are constants. So, we seek an equation of the form $x = a + sv_1 + tw_1$, in which x is expressed in terms of two variables.

If $x \neq 0$, we can solve the given equation $n_1x + n_2y + n_3z = d$ for x in terms of the variables y and z, which gives us

$$x = \frac{d}{n_1} - \frac{n_2}{n_1}y - \frac{n_3}{n_1}z.$$

So, we can simply let $y = s$ and $z = t$ be the parameters. Therefore, the graph of $n_1x + n_2y + n_3z = d$ is the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d/n_1 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} -n_2/n_1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -n_3/n_1 \\ 0 \\ 1 \end{pmatrix}. \tag{12.13}$$

The vectors $\begin{pmatrix} -n_2/n_1 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} -n_3/n_1 \\ 0 \\ 1 \end{pmatrix}$ are linearly independent, since the only way to make the last two components of $s\begin{pmatrix} -n_2/n_1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -n_3/n_1 \\ 0 \\ 1 \end{pmatrix}$ equal 0 is to let $s = t = 0$. Therefore, the graph of (12.13) is a plane.

If $n_1 = 0$, then at least one of n_2 and n_3 is nonzero. We can solve $n_1x+n_2y+n_3z = d$ for the variable corresponding to the nonzero coefficient just as we solved for x above. Then, we can use the other two variables as the parameters, just as we used y and z as parameters above. So, in this case, we can still express the equation in a parametric form matching our algebraic definition of a plane. □

Important: If a, b, c, and d are constants and at least one of a, b, and c is nonzero, then the graph of $ax + by + cz = d$ is a plane.

Problem 12.9: Let $\mathbf{v}$ and $\mathbf{w}$ be linearly independent and let x_0, y_0, and z_0 be constants.

(a) Show that we can always eliminate the parameters from the parametric equation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + s\mathbf{v} + t\mathbf{w}$$

to produce an equation of the form

$$n_1x + n_2y + n_3z = d,$$

where n_1, n_2, n_3, and d are constants.

(b) Show that every point in the graph of the equation of the form $n_1x+n_2y+n_3z = d$ produced by eliminating the parameters is in the graph of the original parametric equation.

Solution for Problem 12.9:

(a) We write the parametric equation as the system

$$\begin{aligned} x &= x_0 + sv_1 + tw_1, \\ y &= y_0 + sv_2 + tw_2, \\ z &= z_0 + sv_3 + tw_3. \end{aligned}$$

We could eliminate the parameters one at a time, but instead, we'll be a little more clever, and try to eliminate them both at once. We think to do this by noticing that the last two terms in each of the equations

above are a dot product:

$$x = x_0 + \begin{pmatrix} s \\ t \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ w_1 \end{pmatrix},$$
$$y = y_0 + \begin{pmatrix} s \\ t \end{pmatrix} \cdot \begin{pmatrix} v_2 \\ w_2 \end{pmatrix},$$
$$z = z_0 + \begin{pmatrix} s \\ t \end{pmatrix} \cdot \begin{pmatrix} v_3 \\ w_3 \end{pmatrix}.$$

This is nice because it allows us to deal with both s and t at the same time—if we eliminate $\begin{pmatrix} s \\ t \end{pmatrix}$ from the system, then we have eliminated both parameters. We now wish to show that there are constants n_1, n_2, and n_3, not all 0, such that we can multiply the first equation by n_1, the second by n_2, and the third by n_3, add the three resulting equations, and thereby eliminate both parameters. After we perform this multiplication and addition, we have

$$\begin{aligned} n_1x + n_2y + n_3z &= n_1x_0 + n_2y_0 + n_3z_0 + n_1 \begin{pmatrix} s \\ t \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ w_1 \end{pmatrix} + n_2 \begin{pmatrix} s \\ t \end{pmatrix} \cdot \begin{pmatrix} v_2 \\ w_2 \end{pmatrix} + n_3 \begin{pmatrix} s \\ t \end{pmatrix} \cdot \begin{pmatrix} v_3 \\ w_3 \end{pmatrix} \\ &= n_1x_0 + n_2y_0 + n_3z_0 + \begin{pmatrix} s \\ t \end{pmatrix} \cdot \left(n_1 \begin{pmatrix} v_1 \\ w_1 \end{pmatrix} + n_2 \begin{pmatrix} v_2 \\ w_2 \end{pmatrix} + n_3 \begin{pmatrix} v_3 \\ w_3 \end{pmatrix} \right). \end{aligned}$$

We can eliminate $\begin{pmatrix} s \\ t \end{pmatrix}$ if we can choose n_1, n_2, n_3, not all 0, such that $n_1 \begin{pmatrix} v_1 \\ w_1 \end{pmatrix} + n_2 \begin{pmatrix} v_2 \\ w_2 \end{pmatrix} + n_3 \begin{pmatrix} v_3 \\ w_3 \end{pmatrix}$ is $\mathbf{0}$. Because any three vectors in two dimensions are linearly dependent, we are guaranteed to be able to find scalars n_1, n_2, n_3, not all 0, such that $n_1 \begin{pmatrix} v_1 \\ w_1 \end{pmatrix} + n_2 \begin{pmatrix} v_2 \\ w_2 \end{pmatrix} + n_3 \begin{pmatrix} v_3 \\ w_3 \end{pmatrix} = \mathbf{0}$. Letting n_1, n_2, n_3 take on these values eliminates the dot product and leaves us

$$n_1x + n_2y + n_3z = n_1x_0 + n_2y_0 + n_3z_0. \tag{12.14}$$

Note that the right side of this equation isn't surprising because the point (x_0, y_0, z_0) must be on the plane.

Letting $d = n_1x_0 + n_2y_0 + n_3z_0$ in Equation (12.14) completes our proof that we can always eliminate the parameters to produce an equation of the form $n_1x + n_2y + n_3z = d$, where n_1, n_2, n_3, and d are constants.

(b) Our work in part (a) shows that we can eliminate the parameters from a parametric equation of the form

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + s\mathbf{v} + t\mathbf{w} \tag{12.15}$$

to produce an equation of the form $n_1x + n_2y + n_3z = d$. This means that every point in the graph of the parametric equation is in the graph of $n_1x + n_2y + n_3z = d$. But we haven't shown that every point in the graph of $n_1x + n_2y + n_3z = d$ is in the graph of the parametric equation, so we don't yet know that the graphs are the same. In Problem 12.8, we showed that we can go from an equation of the form $n_1x + n_2y + n_3z = d$ back to a parametric equation. Specifically, the graph of $n_1x + n_2y + n_3z = d$ is the same as the graph of the parametric equation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + p\mathbf{a} + q\mathbf{b} \tag{12.16}$$

for some point (x_1, y_1, z_1) and some linearly independent vectors $\mathbf{a}$ and $\mathbf{b}$. Now our problem is to show that the graphs of (12.15) and (12.16) are the same.

We know that every point in the graph of (12.15) is in the graph of (12.16). Specifically, we know that (x_0, y_0, z_0) is in the graph of (12.16), so by Problem 12.5, the graph of (12.16) is the same as the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + p\mathbf{a} + q\mathbf{b}. \tag{12.17}$$

Because every point in the graph of (12.15) is in the graph of (12.17), we know that there are two points in (12.17) such that the vector between them is $\mathbf{v}$. In other words $\mathbf{v}$ is in plane (12.17), as is $\mathbf{w}$. Our work in Problem 12.6 then tells us that the plane generated by $\mathbf{a}$ and $\mathbf{b}$ is the same as the plane generated by $\mathbf{v}$ and $\mathbf{w}$. We therefore conclude that the graphs of (12.15) and (12.17) are the same, which means that the graphs of (12.15) and $n_1x + n_2y + n_3z = d$ are the same.

□

At long last, we have proved the following:

Important: Every plane is the graph of an equation of the form $ax + by + cz = d$, where a, b, c, and d are constants.

Make sure you see why this statement is different from, "The graph of the equation $ax + by + cz = d$, where a, b, c, and d are constants, is a plane."

Problem 12.10: In this problem, we find a vector that is orthogonal to every vector in the plane through the origin generated by $\mathbf{v} = 3\mathbf{i} + 4\mathbf{j} - \mathbf{k}$ and $\mathbf{w} = 4\mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$.

(a) Find an equation of the form $n_1x + n_2y + n_3z = 0$ whose graph is the plane through the origin generated by $\mathbf{v}$ and $\mathbf{w}$.

(b) Let $\mathbf{n} = n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}$. Show that $\mathbf{n}$ is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$.

(c) Is $\mathbf{n}$ orthogonal to every vector in the plane?

Solution for Problem 12.10:

(a) The plane is the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\mathbf{v} + t\mathbf{w} = s\begin{pmatrix} 3 \\ 4 \\ -1 \end{pmatrix} + t\begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix}.$$

So, we have

$$\begin{aligned} x &= 3s + 4t, \\ y &= 4s - 2t, \\ z &= -s + 3t. \end{aligned}$$

Adding 3 times the equation for z to the equation for x gives $x + 3z = 13t$. Adding 4 times the equation for z to the equation for y gives $y + 4z = 10t$. Subtracting 13 times $y + 4z = 10t$ from 10 times $x + 3z = 13t$ eliminates t and leaves

$$10x - 13y - 22z = 0.$$

(b) Letting $\mathbf{n} = 10\mathbf{i} - 13\mathbf{j} - 22\mathbf{k}$, we have

$$\begin{aligned} \mathbf{n} \cdot \mathbf{v} &= (10)(3) + (-13)(4) + (-22)(-1) = 0, \\ \mathbf{n} \cdot \mathbf{w} &= (10)(4) + (-13)(-2) + (-22)(3) = 0, \end{aligned}$$

so $\mathbf{n}$ is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$.

(c) Let $\mathbf{u}$ be a vector from the origin to a point in the plane. By definition, such a vector must be a linear combination of $\mathbf{v}$ and $\mathbf{w}$, so we have $\mathbf{u} = s\mathbf{v} + t\mathbf{w}$ for some constants s and t. Therefore, we have

$$\mathbf{n} \cdot \mathbf{u} = \mathbf{n} \cdot (s\mathbf{v} + t\mathbf{w}) = \mathbf{n} \cdot (s\mathbf{v}) + \mathbf{n} \cdot (t\mathbf{w}) = s\mathbf{n} \cdot \mathbf{v} + t\mathbf{n} \cdot \mathbf{w} = 0,$$

which tells us that $\mathbf{n}$ is orthogonal to every vector in the plane through the origin generated by $\mathbf{v}$ and $\mathbf{w}$.

□

The plane in Problem 12.10 passes through the origin. Let's take a look at a plane that doesn't pass through the origin.

Problem 12.11: In Problem 12.7, we found that the plane through the points $(0, 4, 0)$, $(3, -1, -3)$, and $(2, -2, -4)$ is the graph of the equation $x + 3y - 4z = 12$. Let $\mathbf{n} = \mathbf{i} + 3\mathbf{j} - 4\mathbf{k}$, where the components of $\mathbf{n}$ are taken from the coefficients of $x + 3y - 4z = 12$. Is $\mathbf{n}$ orthogonal to the vector between any two points in the plane?

Solution for Problem 12.11: Let (x_1, y_1, z_1) and (x_2, y_2, z_2) be two points in the plane, so

$$x_1 + 3y_1 - 4z_1 = 12,$$
$$x_2 + 3y_2 - 4z_2 = 12.$$

The vector from the first point to the second is $(x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j} + (z_2 - z_1)\mathbf{k}$. These differences appear if we subtract the first equation above from the second equation. This subtraction gives

$$(x_2 - x_1) + 3(y_2 - y_1) - 4(z_2 - z_1) = 0.$$

The expression on the left is $(\mathbf{i} + 3\mathbf{j} - 4\mathbf{k}) \cdot ((x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j} + (z_2 - z_1)\mathbf{k})$. Since this equals 0, the vector $\mathbf{n} = \mathbf{i} + 3\mathbf{j} - 4\mathbf{k}$ is orthogonal to any vector between the two points in the plane. □

We have confirmed that the vector $n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}$ is orthogonal to every vector in the plane $n_1x + n_2y + n_3z = d$ for two different planes. Our solution to Problem 12.11 gives us a guide to prove that this is the case for all planes.

Problem 12.12: Show that the vector $\mathbf{n} = n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}$ is orthogonal to the vector between any two points in the graph of $n_1x + n_2y + n_3z = d$, where n_1, n_2, n_3, and d are constants.

Solution for Problem 12.12: Let (x_1, y_1, z_1) and (x_2, y_2, z_2) be two points in the plane, so

$$n_1x_1 + n_2y_1 + n_3z_1 = d,$$
$$n_1x_2 + n_2y_2 + n_3z_2 = d.$$

Subtracting the first equation from the second gives

$$n_1(x_2 - x_1) + n_2(y_2 - y_1) + n_3(z_2 - z_1) = 0,$$

which we can write as

$$(n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}) \cdot ((x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j} + (z_2 - z_1)\mathbf{k}) = 0.$$

Therefore, $\mathbf{n}$ is orthogonal to any vector between two points in the plane.

Note that this isn't the only vector that is orthogonal to every vector in the plane generated by $\mathbf{v}$ and $\mathbf{w}$. We can multiply the equation above by any scalar k to yield

$$(kn_1\mathbf{i} + kn_2\mathbf{j} + kn_3\mathbf{k}) \cdot ((x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j} + (z_2 - z_1)\mathbf{k}) = 0,$$

which tells us that $k\mathbf{n}$ is orthogonal to any vector in the plane, as well. Specifically, there are two directions that are normal to the plane, the direction of $\mathbf{n}$ and the direction opposite that of $\mathbf{n}$, which is the direction of $-\mathbf{n}$. □

Important: A vector that is orthogonal to all vectors in a certain plane is said to be **normal** to the plane. The vector $a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ is normal to the plane $ax + by + cz = d$.

This should look familiar—we showed back on page 325 that in two dimensions, the vector $a\mathbf{i} + b\mathbf{j}$ is normal to the line $ax + by = c$. We then used that normal vector to find a formula for the distance between a point and a line. Let's see if we can do something similar to find the distance between a point and a plane in three dimensions.

Problem 12.13:

(a) Find an expression for $\text{proj}_{\mathbf{w}}\mathbf{v}$ in terms of $\mathbf{v} \cdot \mathbf{w}$, $\|\mathbf{w}\|$, and $\mathbf{w}$.

(b) The distance between a point and a plane is the length of the shortest segment between the point and a point in the plane. Find an expression for the distance between the point (x_0, y_0, z_0) and the plane $ax + by + cz + d = 0$, where a, b, c, and d are constants.

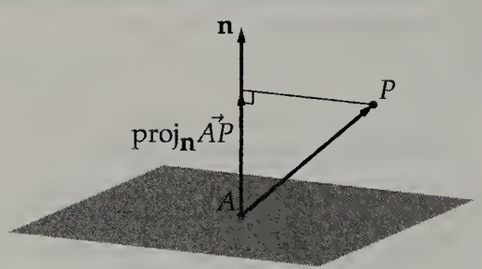
Solution for Problem 12.13:

(a) Flip back to the solution to Problem 9.26 on page 327 and read through it closely, looking for any reference to the fact that the vectors are in two dimensions. You won't find any. The entire proof holds for vectors in three dimensions, as well. So, we don't have to do any more work to see that:

Important: For any vector $\mathbf{v}$ and any nonzero vector $\mathbf{w}$, we have

$$\text{proj}_{\mathbf{w}}\mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{w}\|^2}\mathbf{w}.$$

(b) Looking back at two dimensions saves us a whole lot of trouble on this part, too. Rather than start from scratch, we do exactly what we did in Problem 9.26 in two dimensions. Let point A with coordinates (x, y, z) be on the plane, so that $ax + by + cz + d = 0$. Letting P be (x_0, y_0, z_0), we find the distance from P to the plane by finding the length of the projection of $\vec{AP}$ onto the vector normal to the plane. The vector $\mathbf{n} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ is normal to the plane, so the desired length is



$$\begin{aligned}\left\|\frac{\vec{AP} \cdot \mathbf{n}}{\|\mathbf{n}\|^2}\mathbf{n}\right\| &= \left|\frac{\vec{AP} \cdot \mathbf{n}}{\|\mathbf{n}\|^2}\right| \|\mathbf{n}\| \\ &= \left|\frac{((x_0 - x)\mathbf{i} + (y_0 - y)\mathbf{j} + (z_0 - z)\mathbf{k}) \cdot (a\mathbf{i} + b\mathbf{j} + c\mathbf{k})}{\|\mathbf{n}\|^2}\right| \|\mathbf{n}\| \\ &= \frac{|ax_0 + by_0 + cz_0 - ax - by - cz|}{\|\mathbf{n}\|}.\end{aligned}$$

Since $ax + by + cz + d = 0$, we have $-ax - by - cz = d$. Moreover, $\|\mathbf{n}\| = \sqrt{a^2 + b^2 + c^2}$, so the desired distance between (x_0, y_0, z_0) and $ax + by + cz + d = 0$ is

$$\frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

□

Now that we understand the graph of linear combinations of two vectors, let's take a look at linear combinations of *three* vectors.

Problem 12.14: Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in three dimensions such that $\mathbf{v}$ and $\mathbf{w}$ are linearly independent. We will explore the graph of the parametric equation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = r\mathbf{u} + s\mathbf{v} + t\mathbf{w}. \tag{12.18}$$

(a) Show that if $\mathbf{u}$ is a linear combination of $\mathbf{v}$ and $\mathbf{w}$, then the graph of Equation (12.18) is a plane.

(b) Suppose that $\mathbf{u}$ is not a linear combination of $\mathbf{v}$ and $\mathbf{w}$. Then, what is the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{u} + s\mathbf{v} + t\mathbf{w}?$$

(Notice that the coefficient of $\mathbf{u}$ here is 1, not r.)

(c) Suppose that $\mathbf{u}$ is not a linear combination of $\mathbf{v}$ and $\mathbf{w}$. Then, what is the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2\mathbf{u} + s\mathbf{v} + t\mathbf{w}?$$

How is your answer related to the graph in part (b)?

(d) Suppose that $\mathbf{u}$ is not a linear combination of $\mathbf{v}$ and $\mathbf{w}$. Then, what is the graph of Equation (12.18)?

Solution for Problem 12.14:

(a) If $\mathbf{u}$ is a linear combination of $\mathbf{v}$ and $\mathbf{w}$, then $\mathbf{u} = p\mathbf{v} + q\mathbf{w}$ for some constants p and q. Then, Equation (12.18) becomes

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = r\mathbf{u} + s\mathbf{v} + t\mathbf{w} = r(p\mathbf{v} + q\mathbf{w}) + s\mathbf{v} + t\mathbf{w} = (rp + s)\mathbf{v} + (rq + t)\mathbf{w}.$$

So, the graph consists of all linear combinations of $\mathbf{v}$ and $\mathbf{w}$, which is the plane through the origin generated by $\mathbf{v}$ and $\mathbf{w}$.

(b) Let $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$. Then, the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{u} + s\mathbf{v} + t\mathbf{w} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} + s\mathbf{v} + t\mathbf{w}$$

is the plane through (u_1, u_2, u_3) generated by $\mathbf{v}$ and $\mathbf{w}$.

(c) The graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2\mathbf{u} + s\mathbf{v} + t\mathbf{w} = \begin{pmatrix} 2u_1 \\ 2u_2 \\ 2u_3 \end{pmatrix} + s\mathbf{v} + t\mathbf{w}$$

is the plane through $(2u_1, 2u_2, 2u_3)$ generated by $\mathbf{v}$ and $\mathbf{w}$. As shown in Problem 12.5, this plane is parallel to the plane in part (b) because this plane is generated by the same vectors as the plane in part (b). To show that the planes in these two parts are not the same plane, we will show that it is impossible for the two planes to intersect. If they do intersect, then there must be some point (x, y, z) on both, so there must be

scalars s_1, t_1, s_2, t_2 such that

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} + s_1\mathbf{v} + t_1\mathbf{w} = \begin{pmatrix} 2u_1 \\ 2u_2 \\ 2u_3 \end{pmatrix} + s_2\mathbf{v} + t_2\mathbf{w}.$$

Rearranging this equation gives

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = (s_1 - s_2)\mathbf{v} + (t_1 - t_2)\mathbf{w}.$$

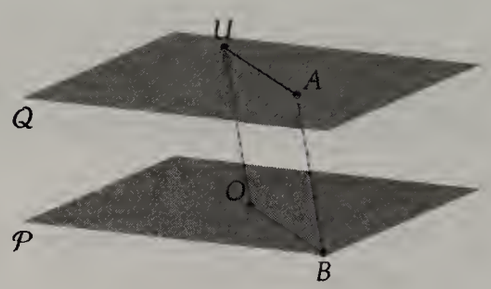
But this equation says that that $\mathbf{u}$ is a linear combination of $\mathbf{v}$ and $\mathbf{w}$, which we know is not the case. Therefore, the plane in part (b) cannot intersect the plane in this part.

(d) Continuing in the vein of part (b) and part (c), for each possible value of r, the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = r\mathbf{u} + s\mathbf{v} + t\mathbf{w}. \tag{12.19}$$

is a plane through (ru_1, ru_2, ru_3) generated by $\mathbf{v}$ and $\mathbf{w}$. Moreover, for the same reason explained in part (c), all of these planes are parallel. So, by varying r, we produce a set of parallel planes. Intuitively, it seems like every point in space is included in one of these planes, but we have to prove it.

We'll start by getting a little more geometric intuition for why each point is in one of these planes. In the diagram at right, plane $\mathcal{P}$ is the plane through the origin O generated by $\mathbf{v}$ and $\mathbf{w}$. Obviously, the graph of (12.19) passes through all points in $\mathcal{P}$, since $\mathcal{P}$ is the graph of (12.19) when $r = 0$.



To show that the graph of (12.19) passes through any point A not in $\mathcal{P}$, we consider the plane Q through A parallel to $\mathcal{P}$. Since $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are linearly independent, when we draw $\mathbf{u}$ from the origin, it is not in $\mathcal{P}$. Therefore, the line through the origin with direction $\mathbf{u}$ intersects Q at some point U, so $\vec{U} = r\mathbf{u}$ for some r. Let B be the point in $\mathcal{P}$ such that $\overleftrightarrow{OB} \parallel \overleftrightarrow{UA}$ and $OB = UA$, so $UOBA$ is a parallelogram (unless U is A, in which case A is obviously in the graph of (12.19)).

Since $\vec{B}$ is in $\mathcal{P}$, we have $\vec{B} = s\mathbf{v} + t\mathbf{w}$ for some s and t. Since $UOBA$ is a parallelogram, we have $\vec{A} = \vec{U} + \vec{B} = r\mathbf{u} + s\mathbf{v} + t\mathbf{w}$, which means A is in the graph of (12.19).

This gives some geometric intuition for why the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = r\mathbf{u} + s\mathbf{v} + t\mathbf{w}. \tag{12.20}$$

is all of $\mathbb{R}^3$. We'll now prove that every point (x, y, z) is on the graph.

Let $ax + by + cz = 0$ be the plane $\mathcal{P}$ through the origin generated by $\mathbf{v}$ and $\mathbf{w}$. Any point (x_0, y_0, z_0) in space is on a plane parallel to $\mathcal{P}$, since for $k_0 = ax_0 + by_0 + cz_0$, the point (x_0, y_0, z_0) is on the graph of $ax + by + cz = k_0$. All we have left is to show that this plane is among the set of parallel planes produced by graphing (12.20). To do so, we just have to show that some multiple of $\mathbf{u}$ produces the plane $ax + by + cz = k_0$. Letting $k_\mathbf{u} = au_1 + bu_2 + cu_3$, the point (u_1, u_2, u_3) is on the graph of $ax + by + cz = k_\mathbf{u}$. Since $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are linearly independent, we know that $\mathbf{u}$ is not in $\mathcal{P}$, so we must have $k_\mathbf{u} \neq 0$. Therefore, we can multiply $au_1 + bu_2 + cu_3 = k_\mathbf{u}$ by $\frac{k_0}{k_\mathbf{u}}$ to give

$$a\left(u_1 \cdot \frac{k_0}{k_\mathbf{u}}\right) + b\left(u_2 \cdot \frac{k_0}{k_\mathbf{u}}\right) + c\left(u_3 \cdot \frac{k_0}{k_\mathbf{u}}\right) = k_0,$$

which means that $\left(u_1 \cdot \frac{k_0}{k_u}, u_2 \cdot \frac{k_0}{k_u}, u_3 \cdot \frac{k_0}{k_u}\right)$ is in the graph of $ax + by + cz = k_0$. This tells us that the graph of $ax + by + cz = k_0$ is the same as the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{k_0}{k_u}\mathbf{u} + s\mathbf{v} + t\mathbf{w},$$

which means (x_0, y_0, z_0) is in the graph of (12.20). Therefore, every point in space is in the graph of (12.20).

□

Problem 12.14 allows us to extend our notion of linear dependence and linear independence to three vectors in three dimensions. We say that $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are **linearly dependent** if and only if there exist constants a_1, a_2, and a_3, not all 0, such that $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 = \mathbf{0}$. Otherwise, we say the vectors are **linearly independent**. Our work in part (d) shows that if $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are linearly independent, then the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = r\mathbf{u} + s\mathbf{v} + t\mathbf{w}$$

is all possible points (x, y, z). More succinctly, we say that this graph is $\mathbb{R}^3$. We also now know that there's nothing to be gained by throwing more vectors into the mix—we can get all of $\mathbb{R}^3$ with linear combinations of three linearly independent vectors, so we can't get "more" by taking linear combinations of more than three vectors.

Problem 12.15: Show that if $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are linearly independent, then for any (x, y, z), there is exactly one ordered triple (r, s, t) such that

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = r\mathbf{u} + s\mathbf{v} + t\mathbf{w}.$$

Solution for Problem 12.15: We know that there is at least one such ordered triple, since the graph of the parametric equation is $\mathbb{R}^3$. Suppose we have both

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = r_1\mathbf{u} + s_1\mathbf{v} + t_1\mathbf{w} \quad \text{and} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = r_2\mathbf{u} + s_2\mathbf{v} + t_2\mathbf{w}.$$

Subtracting the second equation from the first gives

$$\mathbf{0} = (r_1 - r_2)\mathbf{u} + (s_1 - s_2)\mathbf{v} + (t_1 - t_2)\mathbf{w}.$$

Since $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are linearly independent, we must have $r_1 - r_2 = s_1 - s_2 = t_1 - t_2 = 0$, so $r_1 = r_2$, $s_1 = s_2$, and $t_1 = t_2$. Therefore, there cannot be two different ordered triples (r, s, t) that satisfy the equation for the same (x, y, z). Since we have showed that there is at least one ordered triple (r, s, t) that satisfies the problem, but there can't be more than one, we conclude that there is exactly one such ordered triple. □

Exercises

12.2.1 Find a vector with magnitude 1 that is normal to the graph of $4x - 3y + 12z = 5$.

12.2.2 Let $\mathbf{n} = n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}$. Show that if $\mathbf{v}$ is a nonzero vector such that $\mathbf{n} \cdot \mathbf{v} = 0$, then there are two points A and B on the plane $n_1x + n_2y + n_3z = d$ such that $\vec{AB} = \mathbf{v}$.

12.2.3 Find the projection of $3\mathbf{i} + 2\mathbf{j} - 2\mathbf{k}$ onto $4\mathbf{i} - \mathbf{j} + 8\mathbf{k}$.

12.2.4 Find b and c if $\begin{pmatrix} 3 \\ b \\ c \end{pmatrix}$ is normal to the graph of $2x - y + 3x = 5$.

12.2.5 Find all values of b such that the distance between $(3, b, -5)$ and the plane $x + 2y - 2z = 8$ is 8.

12.2.6 We can extend our definition of linear dependence to any number of vectors. Specifically, the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly dependent if and only if there exist constants $a_1, a_2, \ldots, a_n$ such that $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n = \mathbf{0}$ and not all of the a_i are 0. Otherwise, the vectors are linearly independent. Show that any set of four or more vectors in three dimensions is linearly dependent.

12.3 The Cross Product

In Problem 12.10 of the previous section, we found a vector $\mathbf{n}$ that is orthogonal to both $\mathbf{v} = 3\mathbf{i} + 4\mathbf{j} - \mathbf{k}$ and $\mathbf{w} = 4\mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$. We did so by showing that the plane through the origin generated by $\mathbf{v}$ and $\mathbf{w}$ is the graph of $10x - 13y - 22z = 0$. We then showed that the vector $\mathbf{n} = 10\mathbf{i} - 13\mathbf{j} - 22\mathbf{k}$ is orthogonal to any vector between two points in the plane. Specifically, $\mathbf{n}$ is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$.

We start this section by generalizing this process, showing that we can go straight from the linearly independent vectors to the normal vector without bothering with the intermediate step of finding the plane generated by the linearly independent vectors. While doing so, we will discover a vector operation, the cross product, that we use in a great many applications of mathematics to physics and engineering.

Problem 12.16: Find a nonzero vector that is orthogonal to linearly independent vectors $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ and $\mathbf{w} = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k}$, in terms of the components of $\mathbf{v}$ and $\mathbf{w}$.

Solution for Problem 12.16: We can use our solution to Problem 12.10 as a guide. There, found an equation of the form $n_1x + n_2y + n_3z = 0$ whose graph is the plane through the origin generated by $\mathbf{v}$ and $\mathbf{w}$. So, we start with the plane through the origin generated by $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ and $\mathbf{w} = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k}$. Parametric equations whose graph is this plane are

$$\begin{aligned} x &= v_1s + w_1t, \\ y &= v_2s + w_2t, \\ z &= v_3s + w_3t, \end{aligned}$$

where s and t are parameters. We produce an equation of the form $n_1x + n_2y + n_3z = 0$ with the same graph as the parametric equations by eliminating the parameters. We first eliminate s by subtracting v_2 times the first equation from v_1 times the second equation, and by subtracting v_3 times the first equation from v_1 times the third equation. This gives us

$$\begin{aligned} -v_2x + v_1y &= -v_2w_1t + v_1w_2t = (v_1w_2 - v_2w_1)t, \\ -v_3x + v_1z &= -v_3w_1t + v_1w_3t = (v_1w_3 - v_3w_1)t. \end{aligned}$$

Next, we set up elimination of t by multiplying the first equation above by $(v_1w_3 - v_3w_1)$ and the second by $(v_1w_2 - v_2w_1)$:

$$\begin{aligned} -v_2(v_1w_3 - v_3w_1)x + v_1(v_1w_3 - v_3w_1)y &= (v_1w_2 - v_2w_1)(v_1w_3 - v_3w_1)t, \\ -v_3(v_1w_2 - v_2w_1)x + v_1(v_1w_2 - v_2w_1)z &= (v_1w_2 - v_2w_1)(v_1w_3 - v_3w_1)t. \end{aligned}$$

Subtracting the first equation from the second eliminates t and leaves

$$(-v_3(v_1w_2 - v_2w_1) + v_2(v_1w_3 - v_3w_1))\,x - v_1(v_1w_3 - v_3w_1)y + v_1(v_1w_2 - v_2w_1)z = 0.$$

Fortunately, the coefficient of x can be simplified a bit:

$$-v_3(v_1w_2 - v_2w_1) + v_2(v_1w_3 - v_3w_1) = -v_1v_3w_2 + v_2v_3w_1 + v_1v_2w_3 - v_2v_3w_1 = -v_1v_3w_2 + v_1v_2w_3 = v_1(v_2w_3 - v_3w_2).$$

This leaves us

$$v_1(v_2w_3 - v_3w_2)x - v_1(v_1w_3 - v_3w_1)y + v_1(v_1w_2 - v_2w_1)z = 0.$$

If v_1 is nonzero, then we can divide both sides of the equation by v_1 to get

$$(v_2w_3 - v_3w_2)x - (v_1w_3 - v_3w_1)y + (v_1w_2 - v_2w_1)z = 0. \tag{12.21}$$

(If $v_1 = 0$, we can go through the steps above slightly differently to produce this same equation. We eliminated s by subtracting multiples of our parametric equation for x from the parametric equations for y and z. If we started instead by subtracting multiples of the parametric equation for y from the equations for x and z, we'd end up with v_2 times (12.21) rather than v_1 times it.)

Equation (12.21) gives us the desired equation of the form $ax + by + cz = d$ whose graph is the plane through the origin generated by $\mathbf{v}$ and $\mathbf{w}$. We can read the components of a vector that is normal to the plane generated by $\mathbf{v}$ and $\mathbf{w}$ from the coefficients of x, y, and z. Therefore, the vector

$$(v_2w_3 - v_3w_2)\mathbf{i} - (v_1w_3 - v_3w_1)\mathbf{j} + (v_1w_2 - v_2w_1)\mathbf{k}$$

is orthogonal to every vector in the plane generated by $\mathbf{v}$ and $\mathbf{w}$. Specifically, it is orthogonal to $\mathbf{v}$ and $\mathbf{w}$ themselves. □

For any two vectors $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ and $\mathbf{w} = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k}$, we call the vector

$$(v_2w_3 - v_3w_2)\mathbf{i} - (v_1w_3 - v_3w_1)\mathbf{j} + (v_1w_2 - v_2w_1)\mathbf{k}$$

the **cross product** of $\mathbf{v}$ and $\mathbf{w}$, and we denote it as $\mathbf{v} \times \mathbf{w}$. Fortunately, we have an easy way to remember this expression for the cross product. First, we notice that the expressions in parentheses above are determinants of 2×2 matrices, so we can write $\mathbf{v} \times \mathbf{w}$ as

$$\mathbf{v} \times \mathbf{w} = \begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix}\mathbf{k}.$$

The expression on the right side has the same form as one of our expressions for the determinant of a 3×3 matrix. So, we can borrow the notation of the determinant of a 3×3 matrix to remember the expression for the cross product of two vectors:

$$\mathbf{v} \times \mathbf{w} = \begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix}\mathbf{k} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

The vector given by $\mathbf{v} \times \mathbf{w}$ is not the only vector orthogonal to both $\mathbf{v}$ and $\mathbf{w}$. Because $(k\mathbf{p}) \cdot \mathbf{r} = k(\mathbf{p} \cdot \mathbf{r})$ for any scalar k and any vectors $\mathbf{p}$ and $\mathbf{r}$, we can multiply our expression for $\mathbf{v} \times \mathbf{w}$ above by any scalar and produce another vector that has a dot product of 0 with both $\mathbf{v}$ and $\mathbf{w}$. We choose the particular vector above from among all of those that are orthogonal to both $\mathbf{v}$ and $\mathbf{w}$ because it has many convenient properties that we will explore in the next set of problems.

Problems

Problem 12.17: Let $\mathbf{v} = 4\mathbf{i} - \mathbf{j} + 3\mathbf{k}$ and $\mathbf{w} = -\mathbf{i} + 2\mathbf{j} + 7\mathbf{k}$.

(a) Find $\mathbf{v} \times \mathbf{w}$ and $\mathbf{w} \times \mathbf{v}$.

(b) How are the two results you found in part (a) related? For any two vectors $\mathbf{r}$ and $\mathbf{s}$, does the same relationship hold between $\mathbf{r} \times \mathbf{s}$ and $\mathbf{s} \times \mathbf{r}$?

Problem 12.18: Show that for any scalar c and any two vectors $\mathbf{u}$ and $\mathbf{v}$, we have $\mathbf{u} \times (c\mathbf{v}) = (c\mathbf{u}) \times \mathbf{v} = c(\mathbf{u} \times \mathbf{v})$.

Problem 12.19: Show that for any vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in three dimensions, we have $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$ and $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$.

Problem 12.20: Let $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ be nonzero vectors. In this problem, we show that we have $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \, \|\mathbf{v}\| \sin\theta$, where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$.

(a) Find $\|\mathbf{u} \times \mathbf{v}\|$ in terms of the components of $\mathbf{u}$ and $\mathbf{v}$.

(b) What trigonometric function of θ do you already know how to express in terms of $\mathbf{u}$ and $\mathbf{v}$?

(c) Show that $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \, \|\mathbf{v}\| \sin\theta$, where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$.

Problem 12.21: Find $\mathbf{i} \times \mathbf{j}$, $\mathbf{j} \times \mathbf{k}$, and $\mathbf{k} \times \mathbf{i}$.

Problem 12.22:

(a) Find $(6\mathbf{i} + 9\mathbf{j} - 3\mathbf{k}) \times (2\mathbf{i} + 3\mathbf{j} - \mathbf{k})$.

(b) If $\mathbf{r} = c\mathbf{s}$ for some scalar c, then must we have $\mathbf{r} \times \mathbf{s} = \mathbf{0}$?

(c) If $\mathbf{r} \times \mathbf{s} = \mathbf{0}$ and $\mathbf{s} \neq \mathbf{0}$, then must we have $\mathbf{r} = c\mathbf{s}$ for some scalar c?

Problem 12.23: If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are nonzero vectors such that $\mathbf{u} \times \mathbf{v} = \mathbf{u} \times \mathbf{w}$, then must we have $\mathbf{v} = \mathbf{w}$?

Problem 12.24: Let $A = (1,3,2)$, $B = (-3,5,3)$, and $C = (0,-2,-6)$. Let $\mathcal{P}$ be the plane that contains all three of these points. In this problem, we find the equation whose graph is $\mathcal{P}$.

(a) Find a vector that is normal to $\mathcal{P}$.

(b) Find an equation whose graph is $\mathcal{P}$.

Problem 12.25: Let E be $(1,0,-2)$, let F be $(-2,3,-1)$, let G be $(4,2,-3)$, and let H be the point such that $EFGH$ is a parallelogram. Find the area of $EFGH$.

Problem 12.17: Let $\mathbf{v} = 4\mathbf{i} - \mathbf{j} + 3\mathbf{k}$ and $\mathbf{w} = -\mathbf{i} + 2\mathbf{j} + 7\mathbf{k}$.

(a) Find $\mathbf{v} \times \mathbf{w}$ and $\mathbf{w} \times \mathbf{v}$.

(b) How are the two results you found in part (a) related? For any two vectors $\mathbf{r}$ and $\mathbf{s}$, does the same relationship hold between $\mathbf{r} \times \mathbf{s}$ and $\mathbf{s} \times \mathbf{r}$?

Solution for Problem 12.17:

(a) We have

$$\mathbf{v} \times \mathbf{w} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 4 & -1 & 3 \\ -1 & 2 & 7 \end{vmatrix} = \begin{vmatrix} -1 & 3 \\ 2 & 7 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 4 & 3 \\ -1 & 7 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 4 & -1 \\ -1 & 2 \end{vmatrix} \mathbf{k} = -13\mathbf{i} - 31\mathbf{j} + 7\mathbf{k},$$

$$\mathbf{w} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & 2 & 7 \\ 4 & -1 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 7 \\ -1 & 3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} -1 & 7 \\ 4 & 3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} -1 & 2 \\ 4 & -1 \end{vmatrix} \mathbf{k} = 13\mathbf{i} + 31\mathbf{j} - 7\mathbf{k}.$$

(b) In part (a), we see that $\mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v}$. Let's see why this relationship holds for any two vectors. Swapping

two rows of a matrix multiplies the determinant of the matrix by -1, so we have

$$\mathbf{r}\times\mathbf{s} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ r_1 & r_2 & r_3 \\ s_1 & s_2 & s_3 \end{vmatrix} = -\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ s_1 & s_2 & s_3 \\ r_1 & r_2 & r_3 \end{vmatrix} = -\mathbf{s}\times\mathbf{r}.$$

□

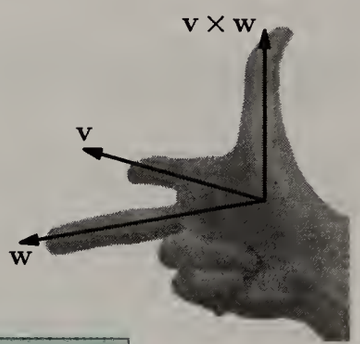
Important: For any two vectors $\mathbf{v}$ and $\mathbf{w}$ in three dimensions, we have

$$\mathbf{v}\times\mathbf{w} = -\mathbf{w}\times\mathbf{v}.$$

WARNING!! The cross product is *not* commutative. If $\mathbf{v}\times\mathbf{w}$ is nonzero, then it does not equal $\mathbf{w}\times\mathbf{v}$.

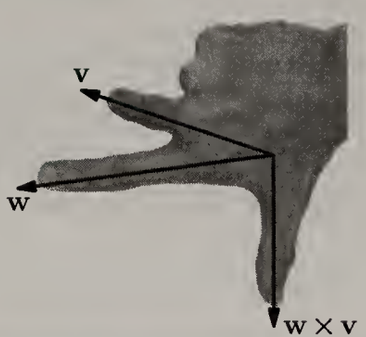
The relationship $\mathbf{v}\times\mathbf{w} = -\mathbf{w}\times\mathbf{v}$ has a geometric significance that makes it easy to visualize. We defined $\mathbf{v}\times\mathbf{w}$ by looking for a vector that is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$. In Problem 12.16, we saw that such a vector is orthogonal to any linear combination of $\mathbf{v}$ and $\mathbf{w}$. That is, the vector $\mathbf{v}\times\mathbf{w}$ is normal to the plane that is generated by $\mathbf{v}$ and $\mathbf{w}$.

However, there are two directions that are normal to this plane, one in each direction away from the plane. The direction of $\mathbf{v}\times\mathbf{w}$ satisfies what we call the **right-hand rule**. To find the direction of $\mathbf{v}\times\mathbf{w}$, start by pointing the index finger of your right hand in the direction of $\mathbf{v}$ and the middle finger of your right hand along $\mathbf{w}$ in such a way that the middle finger is between your index finger and your palm, as shown at right. Then, extend your thumb "up" perpendicular to the plane generated by your index and middle fingers. This is the direction of $\mathbf{v}\times\mathbf{w}$, as shown.



WARNING!! The right-hand rule only works if you are using your right hand, not your left hand. Hence, we have the name "right-hand rule!"

The vector $-\mathbf{v}\times\mathbf{w}$ is in the exact opposite direction from $\mathbf{v}\times\mathbf{w}$. We can visualize this with the right-hand rule. In finding $\mathbf{w}\times\mathbf{v}$, you start with your index finger (of your *right* hand!) fully extended along $\mathbf{w}$ and your middle finger along $\mathbf{v}$ such that your middle finger is between your index finger and your palm. Then, extend your thumb perpendicular to the plane generated by these two fingers, and your thumb will be pointing in exactly the opposite direction from the way it pointed when we found $\mathbf{v}\times\mathbf{w}$, as shown.



Problem 12.18: Show that for any scalar c and any two vectors $\mathbf{u}$ and $\mathbf{v}$, we have $\mathbf{u}\times(c\mathbf{v}) = (c\mathbf{u})\times\mathbf{v} = c(\mathbf{u}\times\mathbf{v})$.

Solution for Problem 12.18: Our expression of the cross product as the determinant of a 3×3 matrix makes this pretty straightforward:

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ cu_1 & cu_2 & cu_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = c\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ cv_1 & cv_2 & cv_3 \end{vmatrix},$$

so

$$(c\mathbf{u}) \times \mathbf{v} = c(\mathbf{u} \times \mathbf{v}) = \mathbf{u} \times (c\mathbf{v}).$$

□

Problem 12.19: Show that $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$ and $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$.

Solution for Problem 12.19: Again, the properties we proved about determinants of 3×3 matrices save the day. We have

$$\mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} + \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 + w_1 & v_2 + w_2 & v_3 + w_3 \end{vmatrix} = \mathbf{u} \times (\mathbf{v} + \mathbf{w}),$$

where we used the principle we proved in Problem 11.21 to combine the sum of determinants of matrices with two rows in common into one determinant.

Combining this result with the fact that $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$ for all $\mathbf{a}$ and $\mathbf{b}$, we have

$$(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = -\mathbf{w} \times (\mathbf{u} + \mathbf{v}) = -\mathbf{w} \times \mathbf{u} - \mathbf{w} \times \mathbf{v} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}.$$

□

Combining Problems 12.18 and 12.19, we see that for any $\mathbf{u}$, both $f(\mathbf{v}) = \mathbf{u} \times \mathbf{v}$ and $f(\mathbf{v}) = \mathbf{v} \times \mathbf{u}$ are linear, so the cross product is bilinear, just like the dot product. (However, as a reminder, the cross product is *not* commutative, but the dot product is.)

Problem 12.20: Show that if θ is the angle between nonzero vectors $\mathbf{u}$ and $\mathbf{v}$, then $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \sin\theta$.

Solution for Problem 12.20: We start by writing an expression for our target, $\|\mathbf{u} \times \mathbf{v}\|$:

$$\begin{aligned} \|\mathbf{u} \times \mathbf{v}\| &= \sqrt{\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}^2 + \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}^2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}^2} \\ &= \sqrt{(u_2v_3 - u_3v_2)^2 + (u_1v_3 - u_3v_1)^2 + (u_1v_2 - u_2v_1)^2} \\ &= \sqrt{u_2^2v_3^2 + u_3^2v_2^2 + u_1^2v_3^2 + u_3^2v_1^2 + u_1^2v_2^2 + u_2^2v_1^2 - 2u_2u_3v_2v_3 - 2u_1v_1u_3v_3 - 2u_1u_2v_1v_2}. \end{aligned}$$

That's quite a mess. However, looking at the other side of the desired equation, $\|\mathbf{u}\| \|\mathbf{v}\| \sin\theta$, we have some hope, since terms like $u_2^2v_3^2$ will appear in the product $\|\mathbf{u}\| \|\mathbf{v}\|$. We don't immediately know how to express $\sin\theta$ in terms of $\mathbf{u}$ and $\mathbf{v}$, but we do know how to deal with $\cos\theta$, since $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos\theta$. Combining this with $\sin^2\theta = 1 - \cos^2\theta$, we can express $\|\mathbf{u}\| \|\mathbf{v}\| \sin\theta$ in terms of the components of $\mathbf{u}$ and $\mathbf{v}$. We start by finding $(\|\mathbf{u}\| \|\mathbf{v}\| \sin\theta)^2$, since this gets rid of the annoying square root signs:

$$\begin{aligned} \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \sin^2\theta &= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 (1 - \cos^2\theta) \\ &= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \left(1 - \frac{(\mathbf{u} \cdot \mathbf{v})^2}{\|\mathbf{u}\|^2 \|\mathbf{v}\|^2}\right) \\ &= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2 \\ &= (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2. \end{aligned}$$

The terms $u_1^2v_1^2$, $u_2^2v_2^2$, and $u_3^2v_3^2$ appear in both the expansion of the product and of the square, so these cancel when we subtract, and we're left with

$$\|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \sin^2\theta = u_2^2v_3^2 + u_3^2v_2^2 + u_1^2v_3^2 + u_3^2v_1^2 + u_1^2v_2^2 + u_2^2v_1^2 - 2u_2u_3v_2v_3 - 2u_1v_1u_3v_3 - 2u_1u_2v_1v_2,$$

which is exactly the square of the expression we found for $\|\mathbf{u} \times \mathbf{v}\|$, so $\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \sin^2\theta$. Since we have $0 \leq \theta \leq \pi$, we know that $\|\mathbf{u}\| \|\mathbf{v}\| \sin\theta$ is nonnegative, so taking the square root of both sides gives us $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \sin\theta$. □

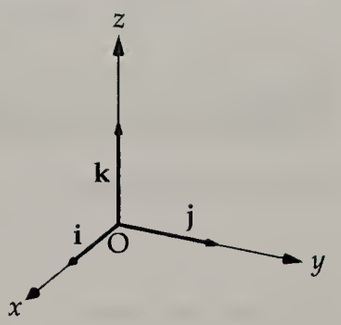
Important: For any nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ in three dimensions, we have

$$\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \, \|\mathbf{v}\| \sin\theta,$$

where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$.

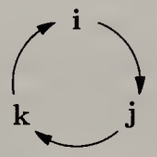
Problem 12.21: Find $\mathbf{i} \times \mathbf{j}$, $\mathbf{j} \times \mathbf{k}$, and $\mathbf{k} \times \mathbf{i}$.

Solution for Problem 12.21: We could use the determinant formula for the cross product to compute the results, but we can also find them with a little geometric intuition. In each case, both vectors have magnitude 1, and the angle between the vectors is $\frac{\pi}{2}$, so the relationship $\|\mathbf{v} \times \mathbf{w}\| = \|\mathbf{v}\| \, \|\mathbf{w}\| \sin\theta$ tells us that the magnitude of each cross product is 1, since $\sin\frac{\pi}{2} = 1$.



All we have left is to determine the direction of each cross product. Each of $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ is orthogonal to the other two. Combining this with the fact that $\mathbf{i} \times \mathbf{j}$ has magnitude 1, we know that $\mathbf{i} \times \mathbf{j}$ is either $\mathbf{k}$ or $-\mathbf{k}$. Applying the right-hand rule reveals that $\mathbf{i} \times \mathbf{j} = \mathbf{k}$. Similarly, we find $\mathbf{j} \times \mathbf{k} = \mathbf{i}$ and $\mathbf{k} \times \mathbf{i} = \mathbf{j}$.

We can visualize these relationships with the diagram at right. For any cross product of the two vectors, if we go in the direction of the arrows (which is in the direction of alphabetical order starting with $\mathbf{i}$) to get from the first vector in the cross product directly to the second, then the cross product equals the third vector. If we must go in the opposite direction of the arrows, such as $\mathbf{k} \times \mathbf{j}$, then the result is the negative of the third vector, so $\mathbf{k} \times \mathbf{j} = -\mathbf{i}$. □



Problem 12.22:

(a) Find $(6\mathbf{i} + 9\mathbf{j} - 3\mathbf{k}) \times (2\mathbf{i} + 3\mathbf{j} - \mathbf{k})$.

(b) If $\mathbf{r} = c\mathbf{s}$ for some scalar c, then must we have $\mathbf{r} \times \mathbf{s} = \mathbf{0}$?

(c) If $\mathbf{r} \times \mathbf{s} = \mathbf{0}$ and $\mathbf{s} \neq \mathbf{0}$, then must we have $\mathbf{r} = c\mathbf{s}$ for some scalar c?

Solution for Problem 12.22:

(a) We have

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 6 & 9 & -3 \\ 2 & 3 & -1 \end{vmatrix} = 0\mathbf{i} + 0\mathbf{j} + 0\mathbf{k} = \mathbf{0}.$$

We also could have seen that this determinant equals $\mathbf{0}$ by noting that the second row is a constant times the third.

(b) Our final observation in part (a) makes it clear that the answer to this part is yes. Letting $\mathbf{s} = s_1\mathbf{i} + s_2\mathbf{j} + s_3\mathbf{k}$, we have

$$\mathbf{r} \times \mathbf{s} = (c\mathbf{s}) \times \mathbf{s} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ cs_1 & cs_2 & cs_3 \\ s_1 & s_2 & s_3 \end{vmatrix} = \mathbf{0},$$

since the second row is a constant times the third. We can also see that $\mathbf{r} \times \mathbf{s} = \mathbf{0}$ if $\mathbf{r} = c\mathbf{s}$ by noting that the angle θ between nonzero vectors $\mathbf{s}$ and $c\mathbf{s}$ is 0 if $c > 0$ and π if $c < 0$. In either case, we have $\sin\theta = 0$, so

$$\|\mathbf{r} \times \mathbf{s}\| = \|\mathbf{r}\| \, \|\mathbf{s}\| \sin\theta = 0.$$

Therefore, we must have $\mathbf{r} \times \mathbf{s} = \mathbf{0}$.

We also could have noted that $(c\mathbf{s}) \times \mathbf{s} = c(\mathbf{s} \times \mathbf{s})$. Since $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$ for any $\mathbf{a}$ and $\mathbf{b}$, letting $\mathbf{a} = \mathbf{b} = \mathbf{s}$ gives $\mathbf{s} \times \mathbf{s} = -\mathbf{s} \times \mathbf{s}$, so $\mathbf{s} \times \mathbf{s} = \mathbf{0}$. Therefore, we have $(c\mathbf{s}) \times \mathbf{s} = c(\mathbf{s} \times \mathbf{s}) = \mathbf{0}$.

(c) Since $\|\mathbf{r} \times \mathbf{s}\| = \|\mathbf{r}\| \|\mathbf{s}\| \sin\theta$, where θ is the angle between $\mathbf{r}$ and $\mathbf{s}$, we must have $\mathbf{r} = \mathbf{0}$ or $\sin\theta = 0$ if $\mathbf{r} \times \mathbf{s} = \mathbf{0}$ and $\mathbf{s} \neq \mathbf{0}$. If $\mathbf{r} = \mathbf{0}$, then $\mathbf{r} = 0\mathbf{s}$. If $\sin\theta = 0$, then $\theta = 0$ or $\theta = \pi$. If $\theta = 0$, then $\mathbf{r}$ and $\mathbf{s}$ are in the same direction, so $\mathbf{r} = c\mathbf{s}$ for some positive c. If $\theta = \pi$, then $\mathbf{r}$ and $\mathbf{s}$ are in opposite directions, so $\mathbf{r} = c\mathbf{s}$ for some negative c. Therefore, in all possible cases, there is a scalar c such that $\mathbf{r} = c\mathbf{s}$.

□

Important: We have $\mathbf{r} \times \mathbf{s} = \mathbf{0}$ if and only if $\mathbf{r}$ and $\mathbf{s}$ are linearly dependent.

Problem 12.23: If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are nonzero vectors such that $\mathbf{u} \times \mathbf{v} = \mathbf{u} \times \mathbf{w}$, then must we have $\mathbf{v} = \mathbf{w}$?

Solution for Problem 12.23: We subtract $\mathbf{u} \times \mathbf{w}$ from both sides and we have

$$\mathbf{u} \times \mathbf{v} - \mathbf{u} \times \mathbf{w} = \mathbf{0}.$$

Applying the properties we proved in Problems 12.18 and 12.19, we have

$$\mathbf{u} \times (\mathbf{v} - \mathbf{w}) = \mathbf{0}.$$

Since there are vectors besides $\mathbf{0}$ whose cross product with $\mathbf{u}$ is $\mathbf{0}$ (namely, $k\mathbf{u}$ for any scalar k), we cannot conclude that $\mathbf{v} - \mathbf{w} = \mathbf{0}$ from the fact that $\mathbf{u} \times (\mathbf{v} - \mathbf{w}) = \mathbf{0}$. Therefore, we cannot conclude that $\mathbf{v} = \mathbf{w}$. □

WARNING!! If nonzero vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ satisfy $\mathbf{u} \times \mathbf{v} = \mathbf{u} \times \mathbf{w}$, then we cannot conclude that $\mathbf{v}$ and $\mathbf{w}$ are equal.

Problem 12.24: Let $A = (1, 3, 2)$, $B = (-3, 5, 3)$, and $C = (0, -2, -6)$. Find an equation whose graph is the plane that passes through all three points.

Solution for Problem 12.24: We know how to find an equation of a plane given a point in the plane and a vector normal to the plane, so all we have to do is find a vector normal to the plane through A, B, and C. We've just learned a tool for finding a vector normal to two other vectors: the cross product. So, to find a vector normal to the plane, we find $\vec{AB} \times \vec{AC}$. We have

$$\vec{AB} = \vec{B} - \vec{A} = (-3 - 1)\mathbf{i} + (5 - 3)\mathbf{j} + (3 - 2)\mathbf{k} = -4\mathbf{i} + 2\mathbf{j} + \mathbf{k},$$
$$\vec{AC} = \vec{C} - \vec{A} = (0 - 1)\mathbf{i} + (-2 - 3)\mathbf{j} + (-6 - 2)\mathbf{k} = -\mathbf{i} - 5\mathbf{j} - 8\mathbf{k}.$$

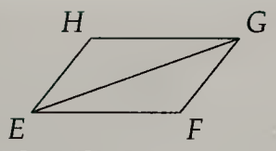
Therefore, we have

$$\vec{AB} \times \vec{AC} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -4 & 2 & 1 \\ -1 & -5 & -8 \end{vmatrix} = -11\mathbf{i} - 33\mathbf{j} + 22\mathbf{k},$$

so $-11\mathbf{i} - 33\mathbf{j} + 22\mathbf{k}$ is normal to the plane through A, B, and C. Factoring out 11, we can write this vector as $11(-\mathbf{i} - 3\mathbf{j} + 2\mathbf{k})$, so $-\mathbf{i} - 3\mathbf{j} + 2\mathbf{k}$ is normal to the plane, and the desired equation is $-x - 3y + 2z = d$ for some constant d. Point A is on the graph of this equation, so we substitute $(1, 3, 2)$ into the equation to find $d = -6$. Therefore, the graph of $-x - 3y + 2z = -6$ is the plane that passes through A, B, and C. □

Problem 12.25: Let E be $(1, 0, -2)$, let F be $(-2, 3, -1)$, let G be $(4, 2, -3)$, and let H be the point such that $EFGH$ is a parallelogram. Find the area of $EFGH$.

Solution for Problem 12.25: The area of a parallelogram equals the length of a side times the distance between that side and the opposite side. We can easily find the side lengths of the parallelogram, but finding the distance between opposite sides is cumbersome. Instead, we note that drawing a diagonal of the parallelogram divides it into two congruent triangles, as shown at right. So, the area of $EFGH$ is twice the area of $\triangle EFG$. But how will we find $[EFG]$? We turn to trigonometry, and we have



$$[EFGH] = 2[EFG] = 2\left(\frac{1}{2}(EF)(FG)\sin\angle EFG\right) = (EF)(FG)\sin\angle EFG.$$

This expression looks familiar! Back in Problem 12.20, we found that $\|\mathbf{u}\times\mathbf{v}\| = \|\mathbf{u}\|\,\|\mathbf{v}\|\sin\theta$, where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$. Applying this here, we have $\left\|\vec{FE}\times\vec{FG}\right\| = (EF)(FG)\sin\angle EFG$, so the area of $\mathcal{A}$ equals $\left\|\vec{FE}\times\vec{FG}\right\|$. Since $\vec{FE} = (1-(-2))\mathbf{i}+(0-3)\mathbf{j}+((-2)-(-1))\mathbf{k} = 3\mathbf{i}-3\mathbf{j}-\mathbf{k}$ and $\vec{FG} = (4-(-2))\mathbf{i}+(2-3)\mathbf{j}+((-3)-(-1))\mathbf{k} = 6\mathbf{i}-\mathbf{j}-2\mathbf{k}$, we have

$$\text{Area of } \mathcal{A} = \left\|\vec{FE}\times\vec{FG}\right\| = \left\|\det\begin{pmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ 3 & -3 & -1\\ 6 & -1 & -2\end{pmatrix}\right\| = \|5\mathbf{i}+0\mathbf{j}+15\mathbf{k}\| = \sqrt{5^2+15^2} = 5\sqrt{10}.$$

□

In the process of solving Problem 12.25, we found another relationship between geometry and the cross product.

Important: The area of a parallelogram with $\mathbf{v}$ and $\mathbf{w}$ as sides is $\|\mathbf{v}\times\mathbf{w}\|$.

Exercises

12.3.1 Let $\mathbf{a} = \begin{pmatrix}3\\4\\-1\end{pmatrix}$, $\mathbf{b} = \begin{pmatrix}-1\\7\\0\end{pmatrix}$, and $\mathbf{c} = \begin{pmatrix}-4\\-1\\3\end{pmatrix}$.

(a) Find $\mathbf{a}\times\mathbf{b}$.

(b) Find $\mathbf{b}\times\mathbf{c}$.

(c) Find $\mathbf{c}\times\mathbf{a}$.

12.3.2 Find an equation whose graph is a plane through $(-1,3,4)$, $(2,1,-5)$, and $(-7,-1,0)$.

12.3.3 Find nonzero vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ such that $\mathbf{u}\times\mathbf{v} = \mathbf{u}\times\mathbf{w}$, but $\mathbf{v}\neq\mathbf{w}$.

12.3.4 Find the area of a triangle with vertices $(-1,3,0)$, $(-2,5,6)$, and $(6,0,-4)$.

12.3.5 Show that $\|\mathbf{u}\times\mathbf{v}\|^2 = \|\mathbf{u}\|^2\,\|\mathbf{v}\|^2 - (\mathbf{u}\cdot\mathbf{v})^2$. This is called the **Lagrange Identity**.

12.3.6★

(a) Show that $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}$.

(b) Show that $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) + \mathbf{b}\times(\mathbf{c}\times\mathbf{a}) + \mathbf{c}\times(\mathbf{a}\times\mathbf{b}) = \mathbf{0}$. This is called the **Jacobi Identity**.

12.4 Geometric Interpretation of the Determinant

In Section 10.5, we found a geometric interpretation of the determinant of a 2×2 matrix. This interpretation described the determinant in terms of area. Moving up to three dimensions, maybe we can describe the determinant of a 3×3 matrix in terms of volume.

In two dimensions, we linked the determinant to the area of a parallelogram. The three-dimensional analogue of a parallelogram is a **parallelepiped**, which is a prism in which all six faces are parallelograms. An example is shown at right.

Problems

Problem 12.26: Let O be the origin, U be (u_1, u_2, u_3), V be (v_1, v_2, v_3), and W be (w_1, w_2, w_3). Also, let $\mathbf{u} = \vec{U}$, $\mathbf{v} = \vec{V}$, and $\mathbf{w} = \vec{W}$. In this problem, we show that the volume of the parallelepiped with $\overline{OU}$, $\overline{OV}$, and $\overline{OW}$ as edges is $|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|$.

(a) Let $\mathcal{P}$ be the parallelogram with sides $\overline{OV}$ and $\overline{OW}$. Find the area of $\mathcal{P}$ in terms of $\mathbf{v}$ and $\mathbf{w}$.

(b) Find the distance from U to the plane of $\mathcal{P}$ in terms of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

(c) Show that the volume of the parallelepiped with $\overline{OU}$, $\overline{OV}$, and $\overline{OW}$ as edges is $|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|$.

Problem 12.27: Let $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$, $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$, and $\mathbf{w} = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k}$. Express $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ as the determinant of a 3×3 matrix whose entries are components of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

Problem 12.28: Find the volume of the tetrahedron with vertices $(3, 4, -1)$, $(-2, 3, 5)$, $(4, -1, 0)$, and $(-5, 5, 2)$.

Problem 12.29: Let $\mathbf{A} = \begin{pmatrix} 3 & -1 & 4 \\ -1 & 1 & -2 \\ -5 & 2 & -7 \end{pmatrix}$.

(a) Find $\det(\mathbf{A})$.

(b) Why does your answer to part (a) tell you that the vectors $\begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix}$, $\begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix}$, and $\begin{pmatrix} -5 \\ 2 \\ -7 \end{pmatrix}$ are not linearly independent?

Problem 12.30: Let $\mathbf{M}$ be a 3×3 matrix.

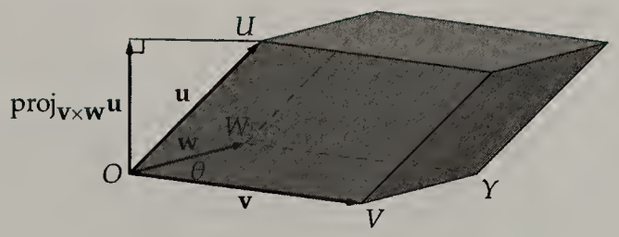
(a) Show that $\det \mathbf{M} = 0$ if and only if the rows of $\mathbf{M}$ are linearly dependent.

(b) Show that $\det \mathbf{M} = 0$ if and only if the columns of $\mathbf{M}$ are linearly dependent.

We found in Section 10.5 that if two given vectors from the same point represent the sides of a parallelogram, then the signed area of the parallelogram equals the determinant of the matrix with these two vectors as columns (or rows). A reasonable guess in three dimensions is that the determinant corresponds to the volume of a parallelepiped. So, let's see if we can find an expression for the volume of a parallelepiped in terms of three vectors that represent edges from the same vertex of the parallelepiped.

Problem 12.26: Let O be the origin, U be (u_1, u_2, u_3), V be (v_1, v_2, v_3), and W be (w_1, w_2, w_3). Also, let $\mathbf{u} = \vec{U}$, $\mathbf{v} = \vec{V}$, and $\mathbf{w} = \vec{W}$. Show that the volume of the parallelepiped with $\overline{OU}$, $\overline{OV}$, and $\overline{OW}$ as edges is $|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|$.

Solution for Problem 12.26: The volume of a parallelepiped equals the area of a face times the distance between that face and the opposite face. Each face of a parallelepiped is a parallelogram, and we know how to find the area of a parallelogram that has two vectors as consecutive sides. We let Y be the point such that $OVYW$ is a parallelogram, and let $\theta = \angle VOW$. Then,



$$[OVYW] = (OV)(OW)\sin\theta = \|\mathbf{v}\times\mathbf{w}\|.$$

To find the distance between $OVYW$ and the opposite face, we find the distance must U to $OVYW$. We discovered how to do this in Problem 12.13; this distance is the magnitude of the projection of $\vec{U}$ onto a vector normal to the plane of $OVYW$. We know how to find such a normal vector: $\mathbf{v}\times\mathbf{w}$ is normal to the plane generated by $\mathbf{v}$ and $\mathbf{w}$. So, since $\vec{U} = \mathbf{u}$, the desired height is $\text{proj}_{\mathbf{v}\times\mathbf{w}}\mathbf{u}$, as shown in the diagram. This gives us

$$\begin{aligned}\text{Volume} = \left\|\text{proj}_{\mathbf{v}\times\mathbf{w}}\mathbf{u}\right\|[OVYW] &= \left\|\frac{\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})}{\|\mathbf{v}\times\mathbf{w}\|^2}(\mathbf{v}\times\mathbf{w})\right\|\|\mathbf{v}\times\mathbf{w}\| \\ &= \frac{|\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})|}{\|\mathbf{v}\times\mathbf{w}\|^2}\|\mathbf{v}\times\mathbf{w}\|^2 \\ &= |\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})|.\end{aligned}$$

□

We can use **signed volume** to get rid of the absolute value sign in $|\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})|$, similar to how we used signed area in two dimensions. Specifically, we say the volume of the parallelepiped with edges $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ is positive if the vector given by $\mathbf{v}\times\mathbf{w}$ makes an acute angle with $\mathbf{u}$, and the volume is negative if $\mathbf{v}\times\mathbf{w}$ makes an obtuse angle with $\mathbf{u}$. Another way to think about this is with the right-hand rule. If $\mathbf{u}$ is on the same side of the plane generated by $\mathbf{v}$ and $\mathbf{w}$ as indicated by the right-hand rule applied to $\mathbf{v}\times\mathbf{w}$, then the volume is positive. If $\mathbf{u}$ is on the opposite side, then the volume is negative.

Important: Let O be the origin, U be (u_1, u_2, u_3), V be (v_1, v_2, v_3), and W be (w_1, w_2, w_3). Also, let $\mathbf{u} = \vec{U}$, $\mathbf{v} = \vec{V}$, and $\mathbf{w} = \vec{W}$. The signed volume of the parallelepiped with $\overline{OU}$, $\overline{OV}$, and $\overline{OW}$ as edges is $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})$.

We call $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})$ a **box product** or **scalar triple product** of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

Now, we're ready to relate the determinant of a 3×3 matrix to volume.

Problem 12.27: Let $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$, $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$, and $\mathbf{w} = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k}$. Express $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})$ as the determinant of a 3×3 matrix whose entries are components of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

Solution for Problem 12.27: We have

$$\begin{aligned}\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = (u_1\mathbf{i}+u_2\mathbf{j}+u_3\mathbf{k})\cdot\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\v_1&v_2&v_3\\w_1&w_2&w_3\end{vmatrix} &= (u_1\mathbf{i}+u_2\mathbf{j}+u_3\mathbf{k})\cdot\left(\begin{vmatrix}v_2&v_3\\w_2&w_3\end{vmatrix}\mathbf{i} - \begin{vmatrix}v_1&v_3\\w_1&w_3\end{vmatrix}\mathbf{j} + \begin{vmatrix}v_1&v_2\\w_1&w_2\end{vmatrix}\mathbf{k}\right) \\ &= u_1\begin{vmatrix}v_2&v_3\\w_2&w_3\end{vmatrix} - u_2\begin{vmatrix}v_1&v_3\\w_1&w_3\end{vmatrix} + u_3\begin{vmatrix}v_1&v_2\\w_1&w_2\end{vmatrix} \\ &= \begin{vmatrix}u_1&u_2&u_3\\v_1&v_2&v_3\\w_1&w_2&w_3\end{vmatrix}.\end{aligned}$$

□

Therefore, we can express the volume of a parallelepiped as a box product, and we can express the box product as a determinant. Putting these together, we have:

Important: Let O be the origin, U be (u_1, u_2, u_3), V be (v_1, v_2, v_3), and W be (w_1, w_2, w_3). The determinant

$$\begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

equals the signed volume of the parallelepiped with $\overline{OU}$, $\overline{OV}$, and $\overline{OW}$ as edges.

We noted on page 367 that if $\mathbf{M}$ is a 2×2 matrix and $\mathbf{v}$ and $\mathbf{w}$ are vectors in 2 dimensions, then the parallelogram with $\mathbf{Mv}$ and $\mathbf{Mw}$ as sides has signed area $\det(\mathbf{M})$ times the signed area of the parallelogram with $\mathbf{v}$ and $\mathbf{w}$. Analogously, if $\mathbf{N}$ is a 3×3 matrix and $\mathbf{r}$, $\mathbf{s}$, and $\mathbf{t}$ are vectors in 3 dimensions, then the parallelepiped with edges $\mathbf{Nr}$, $\mathbf{Ns}$, and $\mathbf{Nt}$ has signed *volume* $\det(\mathbf{N})$ times the signed volume of the parallelepiped with edges $\mathbf{r}$, $\mathbf{s}$, and $\mathbf{t}$.

This observation gives us an intuitive understanding for the following:

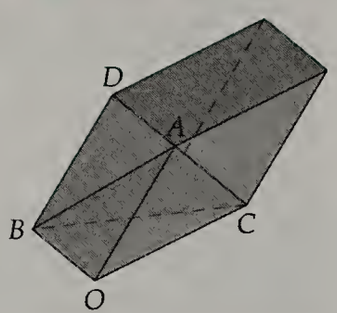
Important: If $\mathbf{A}$ and $\mathbf{B}$ are 3×3 matrices, then

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}).$$

Specifically, if we multiply each of $\mathbf{r}$, $\mathbf{s}$, and $\mathbf{t}$ by $\mathbf{AB}$, we get the same three vectors as if we multiply each first by $\mathbf{B}$ and then multiply the resulting products by $\mathbf{A}$. So, scaling the signed volume in one step by $\det(\mathbf{AB})$ gives the same signed volume as first scaling by $\det(\mathbf{B})$ and then scaling the result by $\det(\mathbf{A})$. Therefore, $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$. (Note that this isn't a proof; it is an intuitive explanation meant to make you have a better feel for the result. Proving that $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$ requires some more serious machinery, or a lot of patience with algebra.)

Problem 12.28: Find the volume of the tetrahedron with vertices $(3, 4, -1)$, $(-2, 3, 5)$, $(4, -1, 0)$, and $(-5, 5, 2)$.

Solution for Problem 12.28: First, we give the points the labels O, A, B, C, respectively. We know how to find the volume of a parallelepiped, so we try to relate the tetrahedron to a parallelepiped we can handle. Specifically, we know how to handle the parallelepiped with $\overline{OA}$, $\overline{OB}$, and $\overline{OC}$ as edges, so we try to relate the desired tetrahedron to this parallelepiped.



Let D be the point such that $AOBD$ is a parallelogram, let $\mathcal{T}$ be tetrahedron $OABC$, let h be the height from C to $AOBD$, and let $\mathcal{P}$ be the parallelepiped with $\overline{OA}$, $\overline{OB}$, and $\overline{OC}$ as edges. Then, we have

$$\text{Volume}(\mathcal{T}) = \frac{h[AOB]}{3} = \frac{h([AOBD]/2)}{3} = \frac{h[AOBD]}{6} = \frac{\text{Volume}(\mathcal{P})}{6}.$$

By Problem 12.27, the volume of $\mathcal{P}$ equals the absolute value of the determinant of the matrix with $\vec{OA}$, $\vec{OB}$, and $\vec{OC}$ as rows. Therefore, we have

$$\text{Volume}(\mathcal{T}) = \frac{1}{6}\left|\det\begin{pmatrix}(-2)-3 & 3-4 & 5-(-1)\\ 4-3 & (-1)-4 & 0-(-1)\\ (-5)-3 & 5-4 & 2-(-1)\end{pmatrix}\right| = \frac{1}{6}\left|\det\begin{pmatrix}-5 & -1 & 6\\ 1 & -5 & 1\\ -8 & 1 & 3\end{pmatrix}\right| = \frac{143}{6}.$$

□

Problem 12.29: Let $\mathbf{A} = \begin{pmatrix} 3 & -1 & 4 \\ -1 & 1 & -2 \\ -5 & 2 & -7 \end{pmatrix}$.

(a) Find $\det(\mathbf{A})$.

(b) Why does your answer to part (a) tell you that the vectors $\begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix}$, $\begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix}$, and $\begin{pmatrix} -5 \\ 2 \\ -7 \end{pmatrix}$ are not linearly independent?

Solution for Problem 12.29:

(a) We have

$$\begin{aligned} \det(\mathbf{A}) &= (3)(1)(-7) + (-1)(-2)(-5) + (4)(-1)(2) - (3)(-2)(2) - (-1)(-1)(-7) - (4)(1)(-5) \\ &= -21 - 10 - 8 - (-12) - (-7) - (-20) = 0. \end{aligned}$$

(b) By Problem 12.27, the determinant equals the signed volume of the parallelepiped with edges represented by the vectors $\mathbf{u} = \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix}$, and $\mathbf{w} = \begin{pmatrix} -5 \\ 2 \\ -7 \end{pmatrix}$. Since the determinant is 0, we know that the volume of this parallelepiped is 0. That is, the distance from the head of $\mathbf{u}$ to the plane generated by $\mathbf{v}$ and $\mathbf{w}$ is 0. This means that $\mathbf{u}$ is in the plane generated by $\mathbf{v}$ and $\mathbf{w}$, so $\mathbf{u}$ is a linear combination of $\mathbf{v}$ and $\mathbf{w}$.

□

Problem 12.30: Let $\mathbf{M}$ be a 3×3 matrix.

(a) Show that $\det \mathbf{M} = 0$ if and only if the rows of $\mathbf{M}$ are linearly dependent.

(b) Show that $\det \mathbf{M} = 0$ if and only if the columns of $\mathbf{M}$ are linearly dependent.

Solution for Problem 12.30:

(a) We have already done all the heavy lifting for this problem. Suppose U, V and W are points in space, and let $\mathbf{M}$ be the matrix whose rows are $\vec{U}$, $\vec{V}$, and $\vec{W}$. Let $\mathcal{P}$ be the parallelepiped with edges $\overline{OU}$, $\overline{OV}$, and $\overline{OW}$. As usual, we let O be the origin.

The volume of $\mathcal{P}$ is 0 if and only if $\vec{U}$, $\vec{V}$, and $\vec{W}$ are in the same plane, which is true if and only if these vectors are linearly dependent. Since $\det(\mathbf{M})$ equals the signed volume of $\mathcal{P}$, we therefore have $\det(\mathbf{M})$ if and only if $\vec{U}$, $\vec{V}$, and $\vec{W}$ are linearly dependent.

Important: The determinant of a 3×3 matrix is 0 if and only if its rows are linearly dependent.

(b) In Problem 11.16, we showed that $\det(\mathbf{M}) = \det(\mathbf{M}^T)$. From part (a), $\det(\mathbf{M}^T) = 0$ if and only if the rows of $\mathbf{M}^T$ are linearly dependent. The rows of $\mathbf{M}^T$ are the columns of $\mathbf{M}$! So, we have $\det(\mathbf{M}) = 0$ if and only if the columns of $\mathbf{M}$ are linearly dependent.

Important: The determinant of a 3×3 matrix is 0 if and only if its columns are linearly dependent.

□

Combining parts (a) and (b) of this solution, we see that:

Important: The rows of a 3×3 matrix are linearly dependent if and only if the columns of the matrix are linearly dependent.

Exercises

12.4.1 The point $(-2, 1, 5)$ is one vertex of a parallelepiped, and it is connected by edges of the parallelepiped to $(4, -1, -2)$, $(-1, 2, -1)$, and $(4, -3, 0)$. What is the volume of the parallelepiped?

12.4.2 Show that $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$.

12.4.3 Suppose that each two rows of matrix $\mathbf{M}$ are orthogonal. Show that the product of the magnitudes of the rows equals $|\det(\mathbf{M})|$.

12.5 Inverse of a 3×3 Matrix

As with 2×2 matrices, an **inverse** of a 3×3 matrix $\mathbf{A}$ is a matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{I}$. We denote the inverse of $\mathbf{A}$ as $\mathbf{A}^{-1}$. In Section 10.6, we showed that a 2×2 matrix has an inverse if and only if the determinant of the matrix is nonzero. In this section, we prove the same condition for the existence of an inverse of a 3×3 matrix.

Problems

Problem 12.31: As a reminder, we say that function f that maps vectors to vectors is a linear function if

(i) $f(a\mathbf{v}) = af(\mathbf{v})$ for all scalars a and vectors $\mathbf{v}$, and

(ii) $f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w})$ for all vectors $\mathbf{v}$ and $\mathbf{w}$.

Show that if g is a linear function that maps vectors to vectors (in three dimensions), then there is a matrix $\mathbf{A}$ such that $g(\mathbf{w}) = \mathbf{Aw}$ for all $\mathbf{w}$. Is this matrix unique for each function g?

Problem 12.32:

(a) Show that if $\det(\mathbf{A}) \neq 0$, then there exists a function f such that for every $\mathbf{c}$, we have $\mathbf{A}(f(\mathbf{c})) = \mathbf{c}$.

(b) Show that the function you found in part (a) is linear.

Problem 12.33:

(a) Show that if $\det(\mathbf{A}) \neq 0$, then there is a matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{I}$.

(b) Is the matrix $\mathbf{B}$ you found in part (a) unique?

(c) Do we also have $\mathbf{AB} = \mathbf{I}$ for the matrix $\mathbf{B}$ you found in part (a)?

Problem 12.34: In this problem, we show that if $\det(\mathbf{A}) = 0$, then $\mathbf{A}$ does not have an inverse.

(a) Show that if $\det(\mathbf{A}) = 0$, then there is some vector $\mathbf{c}$ such that there is no solution $\mathbf{v}$ to $\mathbf{Av} = \mathbf{c}$.

(b) Use part (a) to show that $\mathbf{A}$ does not have an inverse if $\det(\mathbf{A}) = 0$.

Problem 12.31: As a reminder, we say that function f that maps vectors to vectors is a linear function if

(i) $f(a\mathbf{v}) = af(\mathbf{v})$ for all scalars a and vectors $\mathbf{v}$, and

(ii) $f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w})$ for all vectors $\mathbf{v}$ and $\mathbf{w}$.

Show that if g is a linear function that maps vectors to vectors (in three dimensions), then there is a matrix $\mathbf{A}$ such that $g(\mathbf{w}) = \mathbf{A}\mathbf{w}$ for all $\mathbf{w}$. Is this matrix unique for each function g?

Solution for Problem 12.31: Back on page 354, we proved the equivalent statement in two dimensions. We use the same approach here.

We can write any vector $\mathbf{v}$ as $v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$. Applying item (ii) in our definition of a linear function gives

$$g(v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}) = g(v_1\mathbf{i}) + g(v_2\mathbf{j}) + g(v_3\mathbf{k}).$$

Then, applying item (i) to each term on the right, we have

$$g(v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}) = g(v_1\mathbf{i}) + g(v_2\mathbf{j}) + g(v_3\mathbf{k}) = v_1 g(\mathbf{i}) + v_2 g(\mathbf{j}) + v_3 g(\mathbf{k}).$$

Since g maps vectors to vectors, we have

$$g(\mathbf{i}) = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \qquad g(\mathbf{j}) = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \qquad g(\mathbf{k}) = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix},$$

for some scalars $a_{11}, a_{12}, \ldots, a_{33}$. Therefore, we have

$$g\left(\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}\right) = v_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + v_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} + v_3 \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} = \begin{pmatrix} v_1a_{11} + v_2a_{12} + v_3a_{13} \\ v_1a_{21} + v_2a_{22} + v_3a_{23} \\ v_1a_{31} + v_2a_{32} + v_3a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

This tells us that there exists a matrix $\mathbf{A}$ such that $g(\mathbf{v}) = \mathbf{A}\mathbf{v}$ for all vectors $\mathbf{v}$. Moreover, the matrix is unique, since its columns, in order, must be $g(\mathbf{i})$, $g(\mathbf{j})$, and $g(\mathbf{k})$. □

Important: Any linear function g that maps vectors in $\mathbb{R}^3$ to vectors in $\mathbb{R}^3$ can be expressed as multiplication by a unique 3×3 matrix, where the columns of the matrix are $g(\mathbf{i})$, $g(\mathbf{j})$, and $g(\mathbf{k})$.

Problem 12.32:

(a) Show that if $\det(\mathbf{A}) \neq 0$, then there exists a function f such that for every $\mathbf{c}$, we have $\mathbf{A}(f(\mathbf{c})) = \mathbf{c}$.

(b) Show that the function you found in part (a) is linear.

Solution for Problem 12.32:

(a) We showed on page 397 that if $\det(\mathbf{A}) \neq 0$, then the equation $\mathbf{A}\mathbf{v} = \mathbf{c}$ has a unique solution $\mathbf{v}$ for each vector $\mathbf{c}$. Since for each $\mathbf{c}$ there is a unique $\mathbf{v}$ such that $\mathbf{A}\mathbf{v} = \mathbf{c}$, we can define a function $f(\mathbf{c})$ to return the vector $\mathbf{v}$ such that $\mathbf{A}\mathbf{v} = \mathbf{c}$.

(b) We must show the following:

(i) $f(k\mathbf{c}) = kf(\mathbf{c})$ for all scalars k and vectors $\mathbf{c}$, and

(ii) $f(\mathbf{b} + \mathbf{c}) = f(\mathbf{b}) + f(\mathbf{c})$ for all vectors $\mathbf{b}$ and $\mathbf{c}$.

We tackle item (i) first. Let $f(\mathbf{c}) = \mathbf{v}$, so $\mathbf{Av} = \mathbf{c}$. Multiplying both sides by k gives $k\mathbf{Av} = k\mathbf{c}$. Since $k\mathbf{Av} = \mathbf{A}(k\mathbf{v})$, we have $\mathbf{A}(k\mathbf{v}) = k\mathbf{c}$. Therefore, we have $f(k\mathbf{c}) = k\mathbf{v} = kf(\mathbf{c})$.

Turning to item (ii), let $f(\mathbf{b}) = \mathbf{u}$, so $\mathbf{Au} = \mathbf{b}$. Adding $\mathbf{Av} = \mathbf{c}$ to this gives $\mathbf{Au} + \mathbf{Av} = \mathbf{b} + \mathbf{c}$, so $\mathbf{A}(\mathbf{u} + \mathbf{v}) = \mathbf{b} + \mathbf{c}$. Therefore, we have $f(\mathbf{b} + \mathbf{c}) = \mathbf{u} + \mathbf{v} = f(\mathbf{b}) + f(\mathbf{c})$.

Since we have showed that f satisfies conditions (i) and (ii) above, we conclude that f is linear.

□

Problem 12.33:

(a) Show that if $\det(\mathbf{A}) \neq 0$, then there is a matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{I}$.

(b) Is the matrix $\mathbf{B}$ you found in part (a) unique?

(c) Do we also have $\mathbf{AB} = \mathbf{I}$ for the matrix $\mathbf{B}$ you found in part (a)?

Solution for Problem 12.33:

(a) We did all the heavy lifting in Problem 12.32. There, we showed that if $\det(\mathbf{A}) \neq 0$, then there is a linear function f such that, for every $\mathbf{c}$, the vector $f(\mathbf{c})$ is the vector $\mathbf{v}$ such that $\mathbf{Av} = \mathbf{c}$. Because the function is linear, there exists a matrix $\mathbf{B}$ such that $f(\mathbf{c}) = \mathbf{Bc}$. Therefore, if $\mathbf{Av} = \mathbf{c}$, then $\mathbf{Bc} = \mathbf{v}$. Multiplying both sides of $\mathbf{Av} = \mathbf{c}$ on the left by $\mathbf{B}$, we have

$$\mathbf{BAv} = \mathbf{Bc} = \mathbf{v}$$

for all $\mathbf{v}$. Therefore, we must have $\mathbf{BA} = \mathbf{I}$.

(b) The linear function we found in Problem 12.32 is unique for each $\mathbf{A}$, since it is determined by the unique solutions to $\mathbf{Av} = \mathbf{i}$, $\mathbf{Av} = \mathbf{j}$, and $\mathbf{Av} = \mathbf{k}$. In Problem 12.31, we showed that a linear function that maps vectors to vectors can be expressed uniquely as multiplication by a matrix. Therefore, the matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{I}$ is unique.

(c) As explained in part (b), we have $\mathbf{Bc} = \mathbf{v}$ if $\mathbf{Av} = \mathbf{c}$. Multiplying both sides of $\mathbf{Bc} = \mathbf{v}$ on the left by $\mathbf{A}$, we have

$$\mathbf{ABc} = \mathbf{Av} = \mathbf{c}$$

for any vector $\mathbf{c}$, so $\mathbf{AB} = \mathbf{I}$. In other words, if $\mathbf{B}$ is the inverse of $\mathbf{A}$, then $\mathbf{A}$ is the inverse of $\mathbf{B}$.

□

Important: If $\det(\mathbf{A}) \neq 0$, then there is a unique matrix $\mathbf{A}^{-1}$, called the **inverse** of $\mathbf{A}$, such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{AA}^{-1} = \mathbf{I}$.

As with 2×2 matrices, we say that a matrix is **invertible** if it has an inverse. We found that a 2×2 matrix is not invertible if its determinant is 0. Let's see if the same is true of a 3×3 matrix.

Problem 12.34: In this problem, we show that if $\det(\mathbf{A}) = 0$, then $\mathbf{A}$ does not have an inverse.

(a) Show that if $\det(\mathbf{A}) = 0$, then there is some vector $\mathbf{c}$ such that there is no solution $\mathbf{v}$ to $\mathbf{Av} = \mathbf{c}$.

(b) Use part (a) to show that $\mathbf{A}$ does not have an inverse if $\det(\mathbf{A}) = 0$.

Solution for Problem 12.34:

(a) Let the columns of $\mathbf{A}$ be $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$. Since $\det(\mathbf{A}) = 0$, the columns of $\mathbf{A}$ are linearly dependent. So, as explained in Problem 12.14, the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3,$$

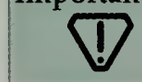
where v_1, v_2, v_3 are parameters, is *not* all of $\mathbb{R}^3$. In other words, there is some vector $\mathbf{c}$ such that there are no values of v_1, v_2, v_3 for which $\mathbf{c} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3$. Since

$$v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3 = v_1\mathbf{A}\mathbf{i} + v_2\mathbf{A}\mathbf{j} + v_3\mathbf{A}\mathbf{k} = \mathbf{A}(v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}),$$

there is no vector $\mathbf{v}$ for which $\mathbf{c} = \mathbf{A}\mathbf{v}$.

(b) If $\mathbf{A}^{-1}$ exists, then we can multiply both sides of $\mathbf{A}\mathbf{v} = \mathbf{c}$ on the left to get $\mathbf{A}^{-1}\mathbf{A}\mathbf{v} = \mathbf{A}^{-1}\mathbf{c}$, so $\mathbf{v} = \mathbf{A}^{-1}\mathbf{c}$. In other words, if $\mathbf{A}$ is invertible, then for any $\mathbf{c}$, the vector $\mathbf{v} = \mathbf{A}^{-1}\mathbf{c}$ is a solution to $\mathbf{A}\mathbf{v} = \mathbf{c}$. But we showed in part (a) that if $\det(\mathbf{A}) = 0$, then there is some vector $\mathbf{c}$ such that there is no solution $\mathbf{v}$ to the equation $\mathbf{A}\mathbf{v} = \mathbf{c}$. So, we conclude that $\mathbf{A}$ is not invertible if $\det(\mathbf{A}) = 0$.

□

Important:  A 3×3 matrix has an inverse if and only if the determinant of the matrix is nonzero.

As a reminder, we say that a 3×3 matrix is **singular** if its determinant is 0 and **nonsingular** if its determinant is nonzero. So, a 3×3 matrix has an inverse if and only if it is nonsingular.

Exercises

12.5.1 If $\mathbf{A}$ is nonsingular, then how are $\det(\mathbf{A})$ and $\det(\mathbf{A}^{-1})$ related?

12.5.2 Find all x such that $\begin{pmatrix} 1 & 5 & 5 \\ -2 & 6 & x \\ x & -2 & 1 \end{pmatrix}$ does not have an inverse.

12.5.3 Must the sum of two singular matrices be singular?

12.5.4 A 3×3 matrix is an **upper triangular matrix** if all the entries below the main diagonal are 0. That is, it is a matrix of the form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}.$$

(a) Show that if all the entries on the main diagonal of an upper triangular matrix (a_{11}, a_{22}, and a_{33} above) are nonzero, then the matrix is invertible.

(b) Show that if an upper triangular matrix is invertible, then its inverse is also an upper triangular matrix.

12.6 Summary

The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are **linearly dependent** if there exist constants $a_1, a_2, \ldots, a_n$, not all 0, such that we have $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n = \mathbf{0}$. Otherwise, we say the vectors are **linearly independent**.

The graph of $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + t\mathbf{v}$ is a line through (x_0, y_0, z_0) with direction vector $\mathbf{v}$.

If $\mathbf{v}$ and $\mathbf{w}$ are linearly independent, then the graph of

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + s\mathbf{v} + t\mathbf{w}$$

is the plane through (x_0, y_0, z_0) generated by $\mathbf{v}$ and $\mathbf{w}$. Any two linearly independent vectors in a plane can be used to generate the plane. A vector is **normal** to a plane if it is orthogonal to every vector in the plane. The graph of any equation of the form $ax + by + cz = d$, where a, b, and c are not all 0, is a plane. Every plane is the graph of an equation of this form, and the vector $a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ is normal to the graph.

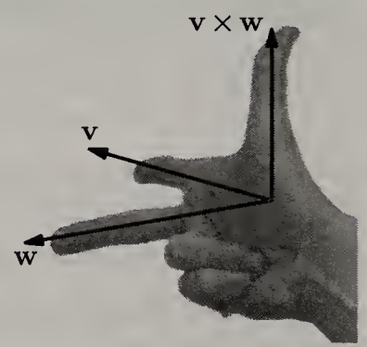
If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are linearly independent, then the graph of $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = r\mathbf{u} + s\mathbf{v} + t\mathbf{w}$ is all of $\mathbb{R}^3$.

We define the **cross product** of two vectors as

$$\mathbf{v} \times \mathbf{w} = \begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix} \mathbf{k} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

The cross product $\mathbf{v} \times \mathbf{w}$ is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$, and has the following properties:

- $\mathbf{v} \times \mathbf{w}$ satisfies the "right-hand rule," which is illustrated at right
- $\mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v}$
- $\mathbf{u} \times (c\mathbf{v}) = (c\mathbf{u}) \times \mathbf{v} = c(\mathbf{u} \times \mathbf{v})$
- $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$ and $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$
- $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \sin\theta$
- $\mathbf{r} \times \mathbf{s} = \mathbf{0}$ if and only if $\mathbf{r}$ and $\mathbf{s}$ are linearly dependent



The signed volume of the parallelepiped with $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ as edges is $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$, where $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ is called a **box product** or **scalar triple product**. We can compute such a product with a determinant:

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

The determinant of a 3×3 matrix also has the following properties:

- If $\mathbf{A}$ and $\mathbf{B}$ are 3×3 matrices, then $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$.
- The determinant of a 3×3 matrix is 0 if and only if its rows are linearly dependent.
- The determinant of a 3×3 matrix is 0 if and only if its columns are linearly dependent.

If $\det(\mathbf{A}) \neq 0$, then there is a unique matrix $\mathbf{A}^{-1}$, called the **inverse** of $\mathbf{A}$, such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. If $\det(\mathbf{A}) = 0$, then $\mathbf{A}$ does not have an inverse.

REVIEW PROBLEMS

12.35 Let $\mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ -5 \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} -3 \\ 4 \\ -2 \end{pmatrix}$, and $\mathbf{p} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

(a) What is the shape of the graph of all points (x, y, z) such that $\mathbf{p} \cdot \mathbf{v} = 4$?

(b) What is the shape of the graph of all points (x, y, z) such that $\mathbf{p} \cdot \mathbf{v} = \mathbf{p} \cdot \mathbf{w}$?

(c) What is the shape of the graph of all points (x, y, z) such that $\mathbf{p} \cdot \mathbf{v} = 4$ and $\mathbf{p} \cdot \mathbf{w} = 4$?

12.36 Find a and b such that $(a, 2, b)$ is on the line through $(2, 4, 2)$ and $(1, 6, -1)$.

12.37 Find a and b if the graphs of $3x - 2y + 8z = 5$ and $ax + by - 4z = 7$ are parallel planes.

12.38 Let ℓ be the line through $(5, 4, 1)$ and $(11, 1, -8)$ and m be the line through $(5, -1, 2)$ and $(0, 14, 7)$.

(a) Show that ℓ and m intersect, and find the point at which they intersect.

(b) Find an equation whose graph is a plane that contains both ℓ and m.

12.39 Let (x_1, y_1, z_1), (x_2, y_2, z_2), and (x_3, y_3, z_3) be three noncollinear points. Show that the graph of

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0$$

is the plane through these three points.

12.40 Let $\mathbf{v} = \begin{pmatrix} 4 \\ -1 \\ 2 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$.

(a) Is $\mathbf{p} = \begin{pmatrix} 9 \\ -4 \\ 1 \end{pmatrix}$ a linear combination of $\mathbf{v}$ and $\mathbf{w}$? If so, find all possible ordered pairs of scalars (a, b) such that $\mathbf{p} = a\mathbf{v} + b\mathbf{w}$.

(b) Suppose $\mathbf{r}$ is a linear combination of $\mathbf{v}$ and $\mathbf{w}$. Show that there is only one ordered pair of scalars (a, b) such that $\mathbf{r} = a\mathbf{v} + b\mathbf{w}$.

12.41

(a) If $\mathbf{u} \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w}$ for all vectors $\mathbf{w}$, then must $\mathbf{u} = \mathbf{v}$?

(b) If $\mathbf{u} \times \mathbf{w} = \mathbf{v} \times \mathbf{w}$ for all vectors $\mathbf{w}$, then must $\mathbf{u} = \mathbf{v}$?

12.42 Find $(3\mathbf{i} + 2\mathbf{j} - \mathbf{k}) \times (-3\mathbf{i} + 4\mathbf{j} + 2\mathbf{k})$.

12.43 If $\mathbf{i} \times \mathbf{v} = \mathbf{k}$, then must $\mathbf{v} = \mathbf{j}$?

12.44 Find an equation of the form $ax + by + cz = d$ whose graph is the plane through the origin generated by $\begin{pmatrix} 4 \\ -1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix}$.

12.45 Find an equation whose graph is the plane that contains both $(-2,3,5)$ and the graph of the parametric equations $x = 3 - 2t$, $y = 4 + t$, $z = -5t$.

12.46 If $\mathbf{a} \times \mathbf{b} = \mathbf{c}$ and $\mathbf{b} \times \mathbf{c} = \mathbf{a}$, then must $\mathbf{c} \times \mathbf{a} = \mathbf{b}$?

12.47

(a) Let $\mathbf{v} = 4\mathbf{i} - \mathbf{j} + 2\mathbf{k}$. Find a matrix $\mathbf{M}$ such that for any vector $\mathbf{w}$, we have $\mathbf{Mw} = \mathbf{v} \times \mathbf{w}$.

(b) Show that for any nonzero vector $\mathbf{u}$, there is a matrix $\mathbf{N}$ such that $\mathbf{Nw} = \mathbf{u} \times \mathbf{w}$ for any vector $\mathbf{w}$.

12.48 Show that if $\mathbf{u} \cdot \mathbf{v} = 0$ and $\mathbf{u} \times \mathbf{v} = \mathbf{0}$, then either $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$.

12.49 Use the geometric interpretation of the determinant to explain why

$$\begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = \begin{vmatrix} u_1 + v_1 & u_2 + v_2 & u_3 + v_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

12.50 Find the inverse of each of the following:

(a) $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

12.51 Suppose that $\mathbf{A}$ and $\mathbf{B}$ both have inverses. Find the inverse of $\mathbf{AB}$ in terms of $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$.

12.52 If $\mathbf{P}$ is the product of an invertible matrix and a non-invertible matrix, then does $\mathbf{P}$ have an inverse?

12.53 Suppose $\mathbf{A}$ is nonsingular. Then, do we have $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$?

Challenge Problems

12.54 We can project a vector on a plane just as we can project it onto a line. Suppose we draw the vector $\mathbf{v}$ from a point O in the plane, where $\mathbf{v}$ is not in the plane. If we let V be the head of $\mathbf{v}$, and let F be a point in the plane such that $\vec{FV}$ is normal to the plane, then the vector $\vec{OF}$ is the projection of the vector onto the plane.

(a) Find a matrix $\mathbf{M}$ such that $\mathbf{Mv}$ is the projection of $\mathbf{v}$ onto the xy-plane (the graph of $z = 0$) for all vectors $\mathbf{v}$.

(b)★ Find a matrix $\mathbf{N}$ such that $\mathbf{Nv}$ is the projection of $\mathbf{v}$ onto the graph of $x + y + z = 0$. **Hints:** 231, 6

12.55 Let $\mathbf{v} = 3\mathbf{i} - 2\mathbf{j} + \mathbf{k}$. Find nonzero vectors $\mathbf{w}$ and $\mathbf{u}$ such that each of $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{u}$ is orthogonal to the other two.

12.56 Let A_{ij} be the minor of row i and column j of matrix $\mathbf{A}$, as described on page 404. Show that if $\mathbf{A}$ is invertible, then

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} A_{11} & -A_{21} & A_{31} \\ -A_{11} & A_{22} & -A_{32} \\ A_{13} & -A_{23} & A_{33} \end{pmatrix}.$$

12.57 Show that if a nonzero vector $\mathbf{v}$ makes an angle of α with $\mathbf{i}$, an angle of β with $\mathbf{j}$, and an angle of γ with $\mathbf{k}$, then we must have $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$. **Hints:** 273

12.58 Show that if $\mathbf{u}$ and $\mathbf{v}$ are not parallel, then the vector $\|\mathbf{v}\|\,\mathbf{u} + \|\mathbf{u}\|\,\mathbf{v}$ bisects the angle between $\mathbf{u}$ and $\mathbf{v}$. **Hints:** 163

12.59 In the text, we developed an expression for the cross product by searching for a vector that is normal to two given vectors. In this problem, we find another way to develop the expression for the cross product of two vectors. So, for this problem, set aside what you already know about the cross product. Suppose we define the cross product such that we have the following:

(i) $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$, and $\mathbf{k} \times \mathbf{i} = \mathbf{j}$

(ii) $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$ for any $\mathbf{a}$ and $\mathbf{b}$

(iii) $\mathbf{a} \times (k\mathbf{b}) = k(\mathbf{a} \times \mathbf{b})$ for any k, $\mathbf{a}$, and $\mathbf{b}$

(iv) $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$ for any $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$

What we have done here is defined the cross product to have some basic algebraic properties that we'd like the cross product to have. In the problems below, you will prove that starting from the properties above, we can derive some of the other properties of the cross product given in the text, as well as the expression for $\mathbf{v} \times \mathbf{w}$ in terms of the components of $\mathbf{v}$ and $\mathbf{w}$.

(a) Show that $\mathbf{v} \times \mathbf{v} = 0$ for any vector $\mathbf{v}$.

(b) Show that if $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ and $\mathbf{w} = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k}$, then $\mathbf{v} \times \mathbf{w} = \begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix}\mathbf{k}$.

(c) Show that $\mathbf{v} \times \mathbf{w}$ is orthogonal to $\mathbf{v}$ and $\mathbf{w}$.

12.60 Show that $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{a} \times \mathbf{c}) = (\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}))\mathbf{a}$. **Hints:** 19

12.61 Show that $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c})$. **Hints:** 36

12.62 Show that if $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ are linearly independent, and each two of them are orthogonal, then for any vector $\mathbf{v}$, we have

$$\mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} + \frac{\mathbf{v} \cdot \mathbf{b}}{\|\mathbf{b}\|^2}\mathbf{b} + \frac{\mathbf{v} \cdot \mathbf{c}}{\|\mathbf{c}\|^2}\mathbf{c}.$$

12.63 Does there exist a 3×2 matrix $\mathbf{M}$ such that for every vector $\mathbf{v}$ in 3 dimensions, there is a vector $\mathbf{u}$ in 2 dimensions such that $\mathbf{Mu} = \mathbf{v}$? **Hints:** 181, 135, 41

12.64 Find the shortest distance from $(0, 1, 2)$ to a point on the line through $(4, 1, -3)$ and $(-2, 5, 1)$. **Hints:** 59

12.65 Let A, B, C, and D be points in space such that the volume of parallelepiped with edges $\overline{AB}$, $\overline{AC}$, and $\overline{AD}$ is 80. **Hints:** 242, 199

(a) If $\overline{AB}$ and $\overline{AC}$ are edges of parallelepiped $\mathcal{P}$ and $\overline{AD}$ is an interior diagonal of $\mathcal{P}$, then what is the volume of $\mathcal{P}$?

(b) If $\overline{AB}$, $\overline{AC}$, and $\overline{AD}$ are face diagonals of $\mathcal{Q}$, then what is the volume of $\mathcal{Q}$?

12.66

(a) Prove that for any vector $\mathbf{v}$, there exists a matrix $\mathbf{M}$ such that $\mathbf{Mw}$ for any $\mathbf{w}$ equals the projection of $\mathbf{w}$ onto $\mathbf{v}$. **Hints:** 210

(b) Suppose $\mathbf{w}$ is a nonzero vector and $\mathbf{M}$ is the matrix such that $\mathbf{Mw} = \text{proj}_{\mathbf{v}}\mathbf{w}$ for all $\mathbf{w}$. Find $\det(\mathbf{M})$. **Hints:** 138

(c) Let $\mathbf{v} = 3\mathbf{i} - 2\mathbf{j} + 5\mathbf{k}$. Find a matrix $\mathbf{M}$ such that $\mathbf{Mw}$ equals the projection of $\mathbf{w}$ onto $\mathbf{v}$ for any vector $\mathbf{w}$.

(d) What is $\mathbf{M}^2$, where $\mathbf{M}$ is your answer to part (a)?

12.67⋆ The distance between two lines in space is the length of the shortest segment that has an endpoint on each line. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be direction vectors of two non-parallel lines in space, k_1 and k_2. Show that for any points A_1 on k_1 and A_2 on k_2, the distance between k_1 and k_2 equals

$$\frac{|\vec{A_1A_2} \cdot (\mathbf{v}_1 \times \mathbf{v}_2)|}{\|\mathbf{v}_1 \times \mathbf{v}_2\|}.$$

Hints: 149, 81, 105

There ain't no rules around here! We're trying to accomplish something! – Thomas Edison

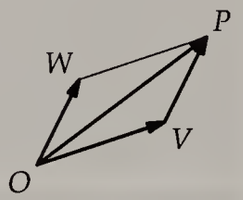
CHAPTER 13 Vector Geometry

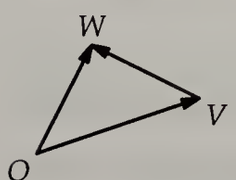
13.1 Introduction

Most of our applications of vectors to geometry start from basic principles that we have already studied. We'll start this section with a reminder of these basics, and a reminder of some of the notation we use when applying vectors to a geometry problems. All of the following relationships apply both to vectors in two dimensions and to vectors in three dimensions.

We depict the vector from a point O to a point V as $\vec{OV}$. When O is the origin, we can express this simply as $\vec{V}$.

We add two vectors using "head-to-tail addition," an example of which is shown at right. Point O is the origin, and we find the vector sum $\vec{V} + \vec{W}$ by drawing a copy of $\vec{W}$ starting at point V, thereby producing $\vec{VP}$ in the diagram. Since $\vec{VP} = \vec{W}$, quadrilateral $OVPW$ is a parallelogram, and we have $\vec{P} = \vec{V} + \vec{W}$.





At left we have a picture of vector subtraction. As explained on page 305, the vector from V to W is $\vec{VW}$, where we have

$$\vec{VW} = \vec{W} - \vec{V}.$$

Since $\vec{W}$ and $\vec{V}$ are vectors from the origin, the relationship $\vec{VW} = \vec{W} - \vec{V}$ gives us a convenient way to express any vector in terms of vectors from the origin.

The last two basic tools we'll use a great deal in this chapter are magnitude and the dot product. The magnitude of $\vec{VW}$ is the length of $\overline{VW}$, so we have

$$\|\vec{VW}\| = VW.$$

The dot product of two vectors is the product of the magnitudes of the two vectors and the cosine of the angle between them. So, if θ is the angle between vectors $\vec{V}$ and $\vec{W}$, then we have

$$\vec{V} \cdot \vec{W} = \vec{W} \cdot \vec{V} = \|\vec{V}\|\|\vec{W}\| \cos\theta.$$

Now, let's get to the problems!

Problems

Problem 13.1: Show that for any three points A, B, and C, we have $\vec{AB} - \vec{AC} = \vec{CB}$.

Problem 13.2: Show that for any distinct points A, B, C, and D, we have $\vec{A} + \vec{C} = \vec{B} + \vec{D}$ if and only if $ABCD$ is a parallelogram.

Problem 13.3:

(a) Let M be the midpoint of $\overline{AB}$. Find $\vec{M}$ in terms of $\vec{A}$ and $\vec{B}$.

(b) Let k be a real number, and let T be on $\overline{AB}$ such that $\frac{AT}{BT} = k$. Find $\vec{T}$ in terms of $\vec{A}$, $\vec{B}$, and k.

Problem 13.4:

(a) Geometrically speaking, what does $\vec{A} \cdot \vec{A}$ equal?

(b) Express AB^2 in terms of $\vec{A} \cdot \vec{A}$, $\vec{A} \cdot \vec{B}$, and $\vec{B} \cdot \vec{B}$. What trigonometric relationship is this known as?

Problem 13.5:

(a) Use the dot product to write an equation in terms of $\vec{A}$, $\vec{B}$, $\vec{C}$, and $\vec{D}$ that is true if and only if $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$.

(b) Use vectors to show that each median of an equilateral triangle is also an altitude of the triangle.

(c) Use vectors to show that if each median of a triangle is also an altitude of the triangle, then the triangle is equilateral.

Problem 13.6:

(a) Show that if A, B, C, and D are points in space such that $AB^2 + CD^2 = BC^2 + AD^2$, then $\vec{AC}$ and $\vec{BD}$ are orthogonal.

(b) Look back at Problem 8.17 on page 274. How is part (a) a more general result than the problem we solved in Problem 8.17? What sorts of problems can we consider with vectors that we cannot tackle with complex numbers?

Problem 13.1: Show that for any three points A, B, and C, we have $\vec{AB} - \vec{AC} = \vec{CB}$.

Solution for Problem 13.1: We tackle this problem by expressing each vector in terms of vectors from the origin, which will allow us to cancel the appearances of A on the left side. But we have to be careful:

WARNING!! We have $\vec{AB} = \vec{B} - \vec{A}$, not $\vec{AB} = \vec{A} - \vec{B}$.

So, we have

$$\vec{AB} = \vec{B} - \vec{A},$$
$$\vec{AC} = \vec{C} - \vec{A}.$$

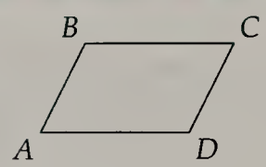
Subtracting the second equation from the first gives

$$\vec{AB} - \vec{AC} = (\vec{B} - \vec{A}) - (\vec{C} - \vec{A}) = \vec{B} - \vec{C} = \vec{CB}.$$

□

Problem 13.2: Show that for any distinct points A, B, C, and D, we have $\vec{A} + \vec{C} = \vec{B} + \vec{D}$ if and only if $ABCD$ is a parallelogram.

Solution for Problem 13.2: We have an "if and only if" problem, so we really have two problems. First, we show that if $\vec{A} + \vec{C} = \vec{B} + \vec{D}$, then $ABCD$ is a parallelogram. We can rearrange the equation as $\vec{B} - \vec{A} = \vec{C} - \vec{D}$, so $\vec{AB} = \vec{DC}$. This tells us that $\overline{AB} \parallel \overline{DC}$ and that $AB = DC$, so A, B, C, and D are vertices of a parallelogram. We know that this parallelogram is $ABCD$, not $BACD$, because $\vec{AB} = \vec{DC}$ tells us that going from A to B is the same direction as going from D to C.



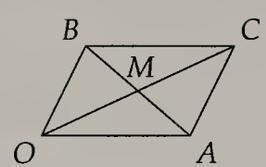
Conversely, if $ABCD$ is a parallelogram, we have $\vec{DC} = \vec{AB}$, so $\vec{C} - \vec{D} = \vec{B} - \vec{A}$. Adding $\vec{A}$ and $\vec{D}$ to both sides gives $\vec{A} + \vec{C} = \vec{B} + \vec{D}$. □

Notice that nothing in our solution to Problem 13.2 assumes that points A, B, C, and D are in the same plane. This means our proof above holds in three dimensions as well as in two dimensions. In other words, it shows that for any points A, B, C, and D in space, we have $\vec{A} + \vec{C} = \vec{B} + \vec{D}$ if and only if $ABCD$ is a parallelogram. It follows that if $\vec{A} + \vec{C} = \vec{B} + \vec{D}$, then A, B, C, and D are coplanar (meaning they are all in the same plane).

Problem 13.3:
(a) Let M be the midpoint of $\overline{AB}$. Find $\vec{M}$ in terms of $\vec{A}$ and $\vec{B}$.
(b) Let k be a real number, and let T be on $\overline{AB}$ such that $\frac{AT}{BT} = k$. Find $\vec{T}$ in terms of $\vec{A}$, $\vec{B}$, and k.

Solution for Problem 13.3:

(a) A natural guess is that $\vec{M}$ is the average of $\vec{A}$ and $\vec{B}$. To see that this guess is correct, let O be the origin, and let C be the point such that $\vec{C} = \vec{A} + \vec{B}$. Then, $OACB$ is a parallelogram. The diagonals of a parallelogram bisect each other, so the midpoint of $\overline{OC}$ is the midpoint of $\overline{AB}$. Therefore, point M is on $\overline{OC}$ such that $OM = \frac{OC}{2}$. This means that $\vec{M}$ has the same direction as $\vec{C}$ but half the magnitude. Therefore, we have



$$\vec{M} = \frac{\vec{C}}{2} = \frac{\vec{A} + \vec{B}}{2}.$$

Important: If M is the midpoint of $\overline{AB}$, then $\vec{M} = (\vec{A} + \vec{B})/2$.

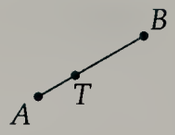
Another way we could have proved this is to note that if M is the midpoint of $\overline{AB}$, then $\vec{AM}$ and $\vec{BM}$ are equal in magnitude, but in opposite directions. We therefore have $\vec{AM} = -\vec{BM}$. Since we also have

$\vec{AM} = \vec{M} - \vec{A}$ and $\vec{BM} = \vec{M} - \vec{B}$, we must have

$$\vec{M} - \vec{A} = -(\vec{M} - \vec{B}).$$

Solving for $\vec{M}$ gives $\vec{M} = (\vec{A} + \vec{B})/2$, as before.

(b) Our second solution to part (a) gives us a pretty good guide to solving this part. Since T is on $\overline{AB}$ such that $AT/BT = k$, we must have $\vec{TA} = k(\vec{BT})$. Notice that we are careful to make sure the two vectors in this equation are pointing in the same direction. Writing $\vec{TA}$ as $\vec{A} - \vec{T}$ and $\vec{BT}$ as $\vec{T} - \vec{B}$, we have

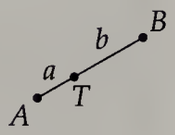


$$\vec{A} - \vec{T} = k(\vec{T} - \vec{B}).$$

Solving for $\vec{T}$ gives

$$\vec{T} = \frac{\vec{A} + k\vec{B}}{1 + k}.$$

One way to think about the relationship among $\vec{T}$, $\vec{A}$, and $\vec{B}$ is that $\vec{T}$ is the weighted average of $\vec{A}$ and $\vec{B}$. This becomes an even more natural way to think about the relationship if we let $a = AT$, and $b = BT$, so $AT/BT = a/b$ and we have



$$\vec{T} = \frac{\vec{A} + \frac{a}{b}\vec{B}}{1 + \frac{a}{b}} = \frac{b\vec{A} + a\vec{B}}{b + a}.$$

It may seem a bit counterintuitive that we weight $\vec{A}$ by the length of $\overline{BT}$ rather than the length of $\overline{AT}$. To get a sense for why this weighting is correct, note that if T is much closer to A than to B, then BT is much larger than AT, so $\frac{b}{b+a}$ is close to 1 and $\frac{a}{b+a}$ is close to 0. This makes $\vec{T}$ close to $\vec{A}$ in our expression above, which is as expected when T is much closer to A than to B.

□

Problem 13.4:

(a) Geometrically speaking, what does $\vec{A} \cdot \vec{A}$ equal?

(b) Express AB^2 in terms of $\vec{A} \cdot \vec{A}$, $\vec{A} \cdot \vec{B}$, and $\vec{B} \cdot \vec{B}$. What trigonometric relationship is this known as?

Solution for Problem 13.4:

(a) Since $\vec{A} \cdot \vec{A} = \|\vec{A}\|^2$, the expression $\vec{A} \cdot \vec{A}$ equals the square of the distance from A to the origin.

(b) Just as $\vec{A} \cdot \vec{A}$ is the square of the distance from A to the origin, the expression $\vec{AB} \cdot \vec{AB}$ equals the square of the distance from A to B. Therefore, we have

$$AB^2 = \vec{AB} \cdot \vec{AB} = (\vec{B} - \vec{A}) \cdot (\vec{B} - \vec{A}) = \vec{B} \cdot \vec{B} - \vec{B} \cdot \vec{A} - \vec{A} \cdot \vec{B} + \vec{A} \cdot \vec{A}.$$

Since $\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A}$, we have

$$AB^2 = \vec{A} \cdot \vec{A} + \vec{B} \cdot \vec{B} - 2\vec{A} \cdot \vec{B}.$$

Letting O be the origin, so that $\|\vec{A}\| = OA$, $\|\vec{B}\| = OB$, and $\angle AOB$ equals the angle between $\vec{A}$ and $\vec{B}$, we recognize the equation above as the Law of Cosines:

$$AB^2 = OA^2 + OB^2 - 2(OA)(OB)\cos\angle AOB.$$

□

Problem 13.5:

(a) Use the dot product to write an equation in terms of $\vec{A}$, $\vec{B}$, $\vec{C}$, and $\vec{D}$ that is true if and only if $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$.

(b) Use vectors to show that each median of an equilateral triangle is also an altitude of the triangle.

(c) Use vectors to show that if each median of a triangle is also an altitude of the triangle, then the triangle is equilateral.

Solution for Problem 13.5:

(a) We have $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$ if and only if the angle between $\vec{AB}$ and $\vec{CD}$ is 90°. An angle between vectors calls for the dot product—the angle between $\vec{AB}$ and $\vec{CD}$ is 90° if and only if $\vec{AB} \cdot \vec{CD} = 0$. Expressing $\vec{AB}$ and $\vec{CD}$ in terms of $\vec{A}$, $\vec{B}$, $\vec{C}$, and $\vec{D}$, we see that $\overleftrightarrow{AB} \perp \overleftrightarrow{CD}$ if and only if

$$(\vec{B} - \vec{A}) \cdot (\vec{D} - \vec{C}) = 0.$$

(b) Let $\triangle ABC$ be an equilateral triangle and let M be the midpoint of $\overline{BC}$. To show that each median of the triangle is also an altitude, we must show that $\overline{AM} \perp \overline{BC}$. From part (a), we have $\overline{AM} \perp \overline{BC}$ if and only if

$$(\vec{M} - \vec{A}) \cdot (\vec{C} - \vec{B}) = 0.$$

Since $\vec{M} = (\vec{B} + \vec{C})/2$, we have

$$(\vec{M} - \vec{A}) \cdot (\vec{C} - \vec{B}) = \left(\frac{\vec{B}}{2} + \frac{\vec{C}}{2} - \vec{A}\right) \cdot (\vec{C} - \vec{B}).$$

That doesn't look like much fun to expand, but we can pull a trick out of our complex number toolbox to help us:

Concept: Choosing your origin wisely can greatly simplify the application of vectors to geometry problems.

Here, letting A be the origin produces a very convenient form for the expression on the right above, and we have

$$\vec{M} \cdot (\vec{C} - \vec{B}) = \left(\frac{\vec{B}}{2} + \frac{\vec{C}}{2}\right) \cdot (\vec{C} - \vec{B}).$$

We factor out $\frac{1}{2}$ and expand the right side. The $\vec{B} \cdot \vec{C}$ term in the expansion cancels with the $-\vec{C} \cdot \vec{B}$ term and we are left with

$$\vec{M} \cdot (\vec{C} - \vec{B}) = \left(\frac{\vec{B}}{2} + \frac{\vec{C}}{2}\right) \cdot (\vec{C} - \vec{B}) = \frac{1}{2}\left(\vec{C} \cdot \vec{C} - \vec{B} \cdot \vec{B}\right).$$

$\vec{B} \cdot \vec{B}$ and $\vec{C} \cdot \vec{C}$ are the squares of the magnitudes of $\vec{B}$ and $\vec{C}$, respectively. Since A is the origin and $\triangle ABC$ is equilateral, the magnitudes of $\vec{B}$ and $\vec{C}$ are the same. So, $\vec{B} \cdot \vec{B} = \vec{C} \cdot \vec{C}$ and we have

$$\vec{M} \cdot (\vec{C} - \vec{B}) = \frac{1}{2}\left(\vec{C} \cdot \vec{C} - \vec{B} \cdot \vec{B}\right) = 0.$$

Since A is the origin, we have $\vec{M} = \vec{AM}$, so $\vec{M} \cdot (\vec{C} - \vec{B}) = 0$ tells us that $\vec{AM} \cdot \vec{BC} = 0$, which means $\overline{AM} \perp \overline{BC}$. Note that our proof only used the fact that $AB = AC$, not the fact that all three sides are equal. So, we have proved that in any isosceles triangle ABC with $AB = AC$, the median from A to $\overline{BC}$ is also an altitude. Applying this fact to each pair of equal sides of an equilateral triangle, we see that all of the medians of an equilateral triangle are also altitudes of the triangle.

(c) We can essentially reverse the steps we took in part (b). If median $\overline{AM}$ is perpendicular to side $\overline{BC}$, then we must have

$$(\vec{M} - \vec{A}) \cdot (\vec{C} - \vec{B}) = 0.$$

We have $\vec{M} = (\vec{B} + \vec{C})/2$, and we let A be the origin, and we have

$$\left(\frac{\vec{B}}{2} + \frac{\vec{C}}{2}\right) \cdot (\vec{C} - \vec{B}) = 0.$$

We multiply by 2 and expand the dot product on the left side. The $\vec{B} \cdot \vec{C}$ and $-\vec{C} \cdot \vec{B}$ terms cancel and we have

$$\vec{C} \cdot \vec{C} - \vec{B} \cdot \vec{B} = 0,$$

which gives us $\vec{C} \cdot \vec{C} = \vec{B} \cdot \vec{B}$. Therefore, the magnitudes of $\vec{B}$ and $\vec{C}$ are the same. This means that B and C are equidistant from the origin. But A is the origin! So, we have $AB = AC$. Similarly, we can show that $AB = BC$ by considering the altitude from B, and we conclude that $\triangle ABC$ is equilateral.

□

Problem 13.6:

(a) Show that if A, B, C, and D are points in space such that $AB^2 + CD^2 = BC^2 + AD^2$, then $\vec{AC}$ and $\vec{BD}$ are orthogonal.

(b) Look back at Problem 8.17 on page 274. How is part (a) a more general result than the problem we solved in Problem 8.17? What sorts of problems can we consider with vectors that we cannot tackle with complex numbers?

Solution for Problem 13.6:

(a) Writing $AB^2 = \vec{AB} \cdot \vec{AB} = (\vec{B} - \vec{A}) \cdot (\vec{B} - \vec{A})$, and likewise for the other terms in the given equation, we have

$$(\vec{B} - \vec{A}) \cdot (\vec{B} - \vec{A}) + (\vec{D} - \vec{C}) \cdot (\vec{D} - \vec{C}) = (\vec{C} - \vec{B}) \cdot (\vec{C} - \vec{B}) + (\vec{D} - \vec{A}) \cdot (\vec{D} - \vec{A}).$$

When we expand the dot products on both sides, $\vec{A} \cdot \vec{A}$, $\vec{B} \cdot \vec{B}$, $\vec{C} \cdot \vec{C}$, and $\vec{D} \cdot \vec{D}$ appear as terms on both sides, so these terms all cancel and we are left with

$$-2\vec{B} \cdot \vec{A} - 2\vec{D} \cdot \vec{C} = -2\vec{B} \cdot \vec{C} - 2\vec{D} \cdot \vec{A}.$$

Dividing by 2 and rearranging this equation gives

$$\vec{B} \cdot \vec{C} - \vec{B} \cdot \vec{A} + \vec{D} \cdot \vec{A} - \vec{D} \cdot \vec{C} = 0.$$

We group the first two terms and the second two terms as

$$\vec{B} \cdot (\vec{C} - \vec{A}) - \vec{D} \cdot (\vec{C} - \vec{A}) = 0.$$

We factor again to get

$$(\vec{B} - \vec{D}) \cdot (\vec{C} - \vec{A}) = 0,$$

so $\vec{DB} \cdot \vec{AC} = 0$. This means that $\vec{AC}$ and $\vec{DB}$ are orthogonal. Since $\vec{BD} = -\vec{DB}$, we see that $\vec{AC} \cdot \vec{BD} = -\vec{AC} \cdot \vec{DB} = 0$, so $\vec{AC}$ and $\vec{BD}$ are orthogonal as well.

(b) The main difference between our result in part (a) and the result in Problem 8.17 is that our work in Problem 8.17 only proves the result in two dimensions. Our work in part (a) shows that it holds in three dimensions, as well. Moreover, it shows that if we define distance in four or more dimensions analogously to our definition of distance in two and three dimensions, then the result holds for higher dimensions, as well. This is an example of one significant advantage vectors have over complex numbers: vectors can be applied to problems in more than two dimensions. □

Exercises

13.1.1

(a) Use vectors to prove that if the diagonals of quadrilateral $ABCD$ bisect each other, then $ABCD$ is a parallelogram.

(b) Use vectors to prove that the midpoints of the sides of any quadrilateral are the vertices of a parallelogram.

13.1.2 We call quadrilateral $ABCD$ is a **kite** if $AB = BC$ and $CD = DA$ (or if $BC = CD$ and $DA = AB$). Use vectors to show that the diagonals of a kite are perpendicular.

13.1.3 Show that if M and N are midpoints of $\overline{AC}$ and $\overline{BD}$, respectively, then $\vec{AB} + \vec{AD} + \vec{CB} + \vec{CD} = 4\vec{MN}$.

13.1.4 A **median** of a quadrilateral connects the midpoint of one side of the quadrilateral to the midpoint of the opposite side of the quadrilateral. Show that the diagonals of a quadrilateral are perpendicular if and only if the medians of the quadrilateral are equal in length.

13.1.5 Show that in any $\triangle ABC$, we have $CA^2 + CB^2 = 2CM^2 + \frac{AB^2}{2}$, where M is the midpoint of $\overline{AB}$.

13.1.6 Let A, B, C, and D be points in space. Show that if $\vec{AB}$ and $\vec{CD}$ are orthogonal and $\vec{AC}$ and $\vec{BD}$ are orthogonal, then $\vec{AD}$ and $\vec{BC}$ are orthogonal.

13.2 Vectors in the Triangle

Problems

Problem 13.7: Show that all three medians of $\triangle ABC$ pass through a point G such that $\vec{G} = (\vec{A} + \vec{B} + \vec{C})/3$.

Problem 13.8: Show that if we let the circumcenter of triangle ABC be the origin, then all three altitudes of $\triangle ABC$ pass through the point H such that $\vec{H} = \vec{A} + \vec{B} + \vec{C}$.

Problem 13.9: The centroid of a triangle is the point at which the triangle's medians are concurrent and the orthocenter of a triangle is the point at which the triangle's altitudes are concurrent. Let G, H, and O be the centroid, orthocenter, and the circumcenter, respectively, of $\triangle ABC$. Show that if O and H are distinct points, then G is on $\overline{OH}$ such that $OG/OH = 1/3$.

Problem 13.10: Let D, E, F be the midpoints of $\overline{BC}$, $\overline{AC}$, and $\overline{AB}$, respectively, and let J, K, L be the feet of the altitudes of $\triangle ABC$ from A, B, C, respectively. Finally, let H be the orthocenter of $\triangle ABC$ and let U, V, W be the midpoints of $\overline{AH}$, $\overline{BH}$, and $\overline{CH}$, respectively. In this problem, we prove that a circle passes through D, E, F, J, K, L, U, V, and W. We call this circle the **nine-point circle** of $\triangle ABC$.

(a) Let the circumcenter of $\triangle ABC$ be the origin. Express $\vec{D}$, $\vec{E}$, and $\vec{F}$ in terms of $\vec{A}$, $\vec{B}$, and $\vec{C}$. Suppose T is the center of the circle that passes through D, E, and F. Use the forms of $\vec{D}$, $\vec{E}$, and $\vec{F}$ to guess $\vec{T}$ in terms of $\vec{A}$, $\vec{B}$, $\vec{C}$. Confirm that your guess is correct by showing that $TD = TE = TF$.

(b) Find $\vec{U}$, $\vec{V}$, and $\vec{W}$ in terms of $\vec{A}$, $\vec{B}$, and $\vec{C}$. Show that $TU = TV = TW$, and that U, V, and W are on the circle from part (a).

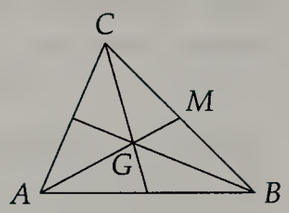
(c) Show that $\overline{UD}$, $\overline{VE}$, and $\overline{WF}$ are diameters of the nine-point circle. Why does this tell us that J, K, and L are on the nine-point circle?

Problem 13.11: Suppose that A', B', and C' are on the circumcircle of $\triangle ABC$ such that $\overline{AA'}$, $\overline{BB'}$, and $\overline{CC'}$ are diameters of the circumcircle. If K, L, and M are the midpoints of $\overline{BC}$, $\overline{AC}$, and $\overline{AB}$, respectively, then show that $\overleftrightarrow{A'K}$, $\overleftrightarrow{B'L}$, and $\overleftrightarrow{C'M}$ are concurrent.

Problem 13.12: Each segment from a vertex of a tetrahedron to the centroid of the opposite face of the tetrahedron is called a **median** of the tetrahedron. Prove that the medians of any tetrahedron are concurrent at a point called the **centroid** of the tetrahedron. Note that the centroid of a triangle cuts each median in a $2:1$ ratio. Can we make a similar statement about the centroid of a tetrahedron?

Problem 13.7: Show that all three medians of $\triangle ABC$ pass through a point G such that $\vec{G} = (\vec{A} + \vec{B} + \vec{C})/3$.

Solution for Problem 13.7: Let M be the midpoint of $\overline{BC}$, so $\vec{M} = (\vec{B} + \vec{C})/2$ and



$$\vec{AM} = \vec{M} - \vec{A} = \frac{\vec{B} + \vec{C}}{2} - \vec{A}.$$

We have

$$\vec{AG} = \vec{G} - \vec{A} = \frac{\vec{A} + \vec{B} + \vec{C}}{3} - \vec{A} = \frac{\vec{B} + \vec{C} - 2\vec{A}}{3} = \frac{2}{3}\left(\frac{\vec{B} + \vec{C}}{2} - \vec{A}\right) = \frac{2}{3}\vec{AM}.$$

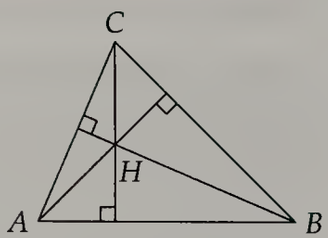
Since $\vec{AG}$ is a scalar multiple of $\vec{AM}$, points A, G, and M are collinear. Moreover, since $\vec{AG} = \frac{2}{3}\vec{AM}$, point G is on $\overline{AM}$ such that $AG = \frac{2}{3}AM$.

There's nothing special in our proof above about vertex A; the same exact steps show that G is on the medians from B and from C, and that the distance from G to each vertex is $\frac{2}{3}$ of the length of the median from that vertex. □

Important: The medians of any triangle are concurrent at a point called the **centroid**. If the centroid of $\triangle ABC$ is G, then $\vec{G} = (\vec{A} + \vec{B} + \vec{C})/3$.

Problem 13.8: Show that if we let the circumcenter of triangle ABC be the origin, then all three altitudes of $\triangle ABC$ pass through the point H such that $\vec{H} = \vec{A} + \vec{B} + \vec{C}$.

Solution for Problem 13.8: Because H is on the altitude from A, we must have $\overleftrightarrow{AH} \perp \overleftrightarrow{BC}$. Vectors give us a nice condition to test if two lines are perpendicular: the dot product. We have $\vec{AH} = \vec{H} - \vec{A} = (\vec{A} + \vec{B} + \vec{C}) - \vec{A} = \vec{B} + \vec{C}$, so



$$\vec{AH} \cdot \vec{BC} = (\vec{B} + \vec{C}) \cdot (\vec{C} - \vec{B}) = \vec{B} \cdot \vec{C} - \vec{B} \cdot \vec{B} + \vec{C} \cdot \vec{C} - \vec{C} \cdot \vec{B}.$$

Since $\vec{B} \cdot \vec{C} = \vec{C} \cdot \vec{B}$, two terms cancel, leaving

$$\vec{AH} \cdot \vec{BC} = \vec{C} \cdot \vec{C} - \vec{B} \cdot \vec{B} = \|\vec{C}\|^2 - \|\vec{B}\|^2.$$

We'd like to show that this expression equals 0, but now we seem stuck.

Concept: When stuck on a problem, look back at the problem for any information you haven't used.

We haven't used the fact that the circumcenter of $\triangle ABC$ is the origin. The circumcenter of a triangle is equidistant from its vertices, so $\|\vec{A}\| = \|\vec{B}\| = \|\vec{C}\|$. Therefore, we have $\vec{AH} \cdot \vec{BC} = \|\vec{C}\|^2 - \|\vec{B}\|^2 = 0$, which means that $\overleftrightarrow{AH} \perp \overleftrightarrow{BC}$. Similarly, we can show that $\vec{BH} \cdot \vec{AC} = \vec{CH} \cdot \vec{AB} = 0$, so H is on all three altitudes of $\triangle ABC$. □

Important: The altitudes of any triangle are concurrent at a point called the **orthocenter**. Let H be the orthocenter of $\triangle ABC$. If the circumcenter of $\triangle ABC$ is the origin, then $\vec{H} = \vec{A} + \vec{B} + \vec{C}$.

WARNING!! Let H be the orthocenter of $\triangle ABC$. Our proof above only shows that we have $\vec{H} = \vec{A} + \vec{B} + \vec{C}$ if the circumcenter of $\triangle ABC$ is the origin.

We've already seen the power of choosing the origin wisely several times in problems involving the application of complex numbers or vectors to geometry problems. Problem 13.8 shows us a frequently useful choice for problems involving triangles.

Concept: When applying vectors to a problem involving a triangle, choosing the circumcenter as the origin can greatly simplify the problem.

As we saw in Problem 13.8, this tactic is useful because the magnitudes of the vectors from the origin to the vertices of the triangle are all equal if we let the circumcenter of the triangle be the origin. The simple form we found for the orthocenter when we let the circumcenter be the origin gives us another good reason to make this choice of origin in problems involving the orthocenter.

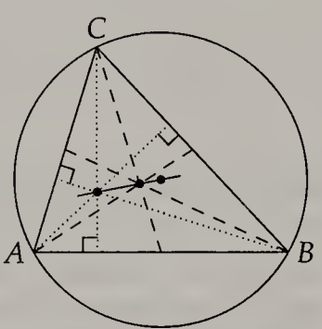
Problem 13.9: Let G, H, and O be the centroid, orthocenter, and the circumcenter of $\triangle ABC$. Show that if H and O are distinct points, then G is on $\overline{OH}$ such that $OG/OH = 1/3$.

Solution for Problem 13.9: We have a convenient vector representation for the centroid, and if we let the circumcenter be the origin, we have a very convenient representation for the circumcenter and a nice one for the orthocenter. Letting O be the origin, we have $\vec{OG} = \vec{G}$ and $\vec{OH} = \vec{H}$. Using our nice expressions for $\vec{G}$ and $\vec{H}$, we have

$$\vec{G} = \frac{\vec{A} + \vec{B} + \vec{C}}{3} = \frac{\vec{H}}{3}.$$

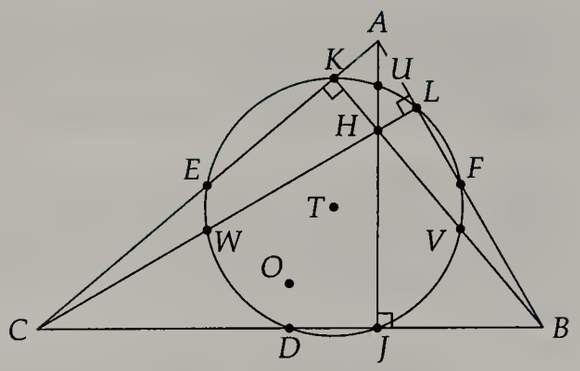
Therefore, G is on $\overline{OH}$ such that $OG/OH = 1/3$. □

The line through the orthocenter, centroid, and circumcenter of a non-equilateral triangle is called the **Euler line** of the triangle. (If a triangle is equilateral, these three points are all the same.) In the diagram at right, the altitudes are dotted, the medians are dashed, and the solid line passing through the orthocenter, centroid, and circumcenter of the triangle is the Euler line.



Problem 13.10: Let D, E, F be the midpoints of $\overline{BC}$, $\overline{AC}$, and $\overline{AB}$, respectively, and let J, K, L be the feet of the altitudes of $\triangle ABC$ from A, B, C, respectively. Finally, let H be the orthocenter of $\triangle ABC$ and let U, V, W be the midpoints of $\overline{AH}$, $\overline{BH}$, and $\overline{CH}$, respectively. Show that a single circle passes through D, E, F, J, K, L, U, V, and W. This circle is called the **nine-point circle** of $\triangle ABC$.

Solution for Problem 13.10: We can show that all nine points are on the same circle by finding a single point that is the same distance from all nine points. We think to attack this problem with vectors because we can easily handle midpoints and the orthocenter with vectors. But to deal with the orthocenter, we must remember to let the circumcenter of $\triangle ABC$ be the origin.



We start with the easiest points to work with, the midpoints of the sides, since these are the simplest to express with vectors. We have

$$\vec{D} = \frac{\vec{B}+\vec{C}}{2},$$
$$\vec{E} = \frac{\vec{A}+\vec{C}}{2},$$
$$\vec{F} = \frac{\vec{A}+\vec{B}}{2}.$$

The desired circle must pass through all three of these points. We let the circle have center T, and will refer to the circle as $\odot T$. Can we use the expressions for $\vec{D}$, $\vec{E}$, and $\vec{F}$ to deduce $\vec{T}$ in terms of $\vec{A}$, $\vec{B}$, and $\vec{C}$?

Concept: Don't overlook "guess-and-check" as a strategy, even in advanced problems. Guessing and then confirming that your guess is correct has a long and glorious history in mathematics and science.

We seek a point T such that $\vec{DT}$, $\vec{ET}$, and $\vec{FT}$ all have the same magnitude. Since $\vec{DT} = \vec{T} - \vec{D}$, $\vec{ET} = \vec{T} - \vec{E}$, and $\vec{FT} = \vec{T} - \vec{F}$, a natural choice for $\vec{T}$ that offers a lot of cancellation in these differences is

$$\vec{T} = \frac{\vec{A}+\vec{B}+\vec{C}}{2}.$$

This expression gives very simple results for $\vec{DT}$, $\vec{ET}$, and $\vec{FT}$:

$$\vec{DT} = \vec{T} - \vec{D} = \frac{\vec{A}+\vec{B}+\vec{C}}{2} - \frac{\vec{B}+\vec{C}}{2} = \frac{\vec{A}}{2},$$
$$\vec{ET} = \vec{T} - \vec{E} = \frac{\vec{A}+\vec{B}+\vec{C}}{2} - \frac{\vec{A}+\vec{C}}{2} = \frac{\vec{B}}{2},$$
$$\vec{FT} = \vec{T} - \vec{F} = \frac{\vec{A}+\vec{B}+\vec{C}}{2} - \frac{\vec{A}+\vec{B}}{2} = \frac{\vec{C}}{2}.$$

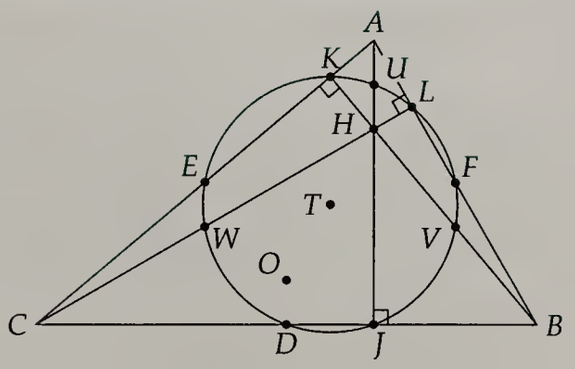
Recalling that the origin is the circumcenter of $\triangle ABC$, we see that this choice of $\vec{T}$ is a winner, since $\|\vec{A}\| = \|\vec{B}\| = \|\vec{C}\|$ gives us $\|\vec{DT}\| = \|\vec{ET}\| = \|\vec{FT}\|$. So, T is the center of the circle that passes through D, E, and F; we'll refer to this circle as $\odot T$. Moreover, because we know that $\|\vec{A}\| = R$, the circumradius of $\triangle ABC$, and $\|\vec{DT}\| = \|\vec{A}\|/2$, we know that the radius of the circle through D, E, and F is $R/2$. We also recognize that

$$\vec{T} = \frac{\vec{A}+\vec{B}+\vec{C}}{2} = \frac{\vec{H}}{2},$$

so T is the midpoint of the segment from the circumcenter to the orthocenter. This means T is on the Euler line, too!

Next, we turn to U, V, and W, because we know how to handle the orthocenter with vectors. Since U, V, and W are the midpoints of $\overline{AH}$, $\overline{BH}$, and $\overline{CH}$, respectively, we have

$$\vec{U} = \frac{\vec{A}+\vec{H}}{2} = \frac{2\vec{A}+\vec{B}+\vec{C}}{2},$$
$$\vec{V} = \frac{\vec{B}+\vec{H}}{2} = \frac{\vec{A}+2\vec{B}+\vec{C}}{2},$$
$$\vec{W} = \frac{\vec{C}+\vec{H}}{2} = \frac{\vec{A}+\vec{B}+2\vec{C}}{2}.$$



We need to show that each of U, V, and W is $R/2$ from T, so we find the magnitudes of $\vec{UT}$, $\vec{VT}$, and $\vec{WT}$:

$$\|\vec{UT}\| = \|\vec{T}-\vec{U}\| = \left\|\frac{\vec{A}+\vec{B}+\vec{C}}{2} - \frac{2\vec{A}+\vec{B}+\vec{C}}{2}\right\| = \left\|-\frac{\vec{A}}{2}\right\| = \left\|\frac{\vec{A}}{2}\right\| = \frac{R}{2},$$
$$\|\vec{VT}\| = \|\vec{T}-\vec{V}\| = \left\|\frac{\vec{A}+\vec{B}+\vec{C}}{2} - \frac{\vec{A}+2\vec{B}+\vec{C}}{2}\right\| = \left\|-\frac{\vec{B}}{2}\right\| = \left\|\frac{\vec{B}}{2}\right\| = \frac{R}{2},$$
$$\|\vec{WT}\| = \|\vec{T}-\vec{W}\| = \left\|\frac{\vec{A}+\vec{B}+\vec{C}}{2} - \frac{\vec{A}+\vec{B}+2\vec{C}}{2}\right\| = \left\|-\frac{\vec{C}}{2}\right\| = \left\|\frac{\vec{C}}{2}\right\| = \frac{R}{2}.$$

So, $\odot T$ passes through U, V, and W in addition to D, E, and F. Six points down, three to go!

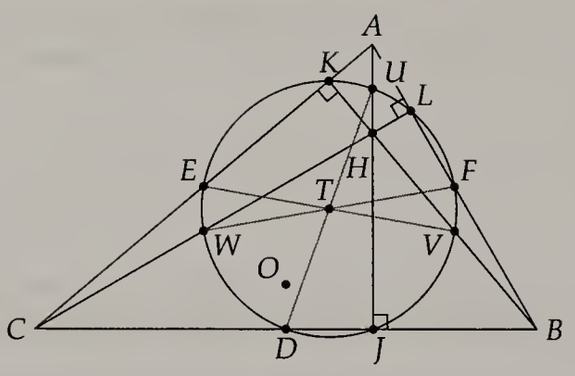
The last three points, the feet of the altitudes of $\triangle ABC$, look a lot more challenging. We haven't seen simple vector expressions in terms of $\vec{A}$, $\vec{B}$, and $\vec{C}$ for the feet of the altitudes of $\triangle ABC$. So, let's try learning a little bit more about the six points that we have already proved are on the circle. Looking closely at our diagram, it looks like $\overline{DU}$, $\overline{EV}$, and $\overline{FW}$ each pass through T. If they do, they're diameters of $\odot T$.

Concept: Precise diagrams can help you discover important relationships in geometry problems.

Of course, it's not good enough that $\overline{DU}$, $\overline{EV}$, and $\overline{FW}$ *look like* diameters. We have to prove it. Vectors can help us with that, too. The midpoint of $\overline{DU}$ is

$$\frac{\vec{D}+\vec{U}}{2} = \frac{\frac{\vec{B}+\vec{C}}{2} + \frac{2\vec{A}+\vec{B}+\vec{C}}{2}}{2} = \frac{\vec{A}+\vec{B}+\vec{C}}{2} = \vec{T}.$$

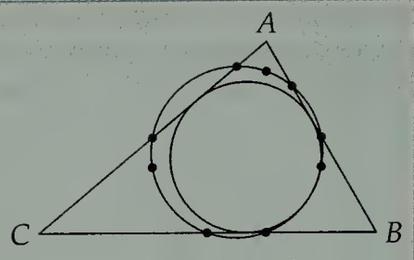
Success! T is the midpoint of $\overline{DU}$, so $\overline{DU}$ is indeed a diameter of $\odot T$. Similarly, $\frac{\vec{E}+\vec{V}}{2} = \vec{T}$ and $\frac{\vec{F}+\vec{W}}{2} = \vec{T}$, so $\overline{EV}$ and $\overline{FW}$ are diameters of $\odot T$ as well.



Now, we're ready to tackle J, K, and L. The circumcircle of a right triangle is the circle with the hypotenuse of the triangle as a diameter. Therefore, the circumcircle of right triangle DJU is the circle with hypotenuse $\overline{DU}$ as diameter. But this is $\odot T$! Since $\odot T$ is the circumcircle of $\triangle DJU$, it passes through J. Similarly, $\odot T$ is the circumcircle of right triangles EKV and FLW, so K and L are on $\odot T$, as well. □

Extra! The midpoint of any segment from the orthocenter of a triangle to a point on the circumcircle of the triangle is on the nine-point circle of the triangle.

Sidenote:

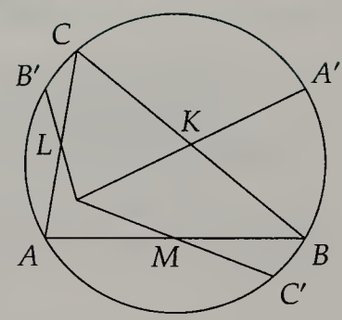


The nine-point circle has interesting relationships to other parts of a triangle. For example, the incircle of a triangle is tangent to the nine-point circle, as shown at right (except when the triangle is equilateral, in which case the two circles are the same). The nine-point circle is also tangent to the **excircles** of $\triangle ABC$, which are the three circles outside $\triangle ABC$ that are tangent to one side of the triangle and tangent to the extensions of the other two sides. That the nine-point circle of a triangle is tangent to the incircle and the excircles of the triangle is sometimes referred to as **Feuerbach's Theorem**.

One direct result of Feuerbach's Theorem is that the inradius of a triangle is no more than half the circumradius of the triangle. See if you can figure out why!

Problem 13.11: Suppose that A', B', and C' are on the circumcircle of $\triangle ABC$ such that $\overline{AA'}$, $\overline{BB'}$, and $\overline{CC'}$ are diameters of the circumcircle. If K, L, and M are the midpoints of $\overline{BC}$, $\overline{AC}$, and $\overline{AB}$, respectively, then show that $\overleftrightarrow{A'K}$, $\overleftrightarrow{B'L}$, and $\overleftrightarrow{C'M}$ are concurrent.

Solution for Problem 13.11: We use vectors because we know how to express all the key points with vectors in terms of $\vec{A}$, $\vec{B}$, and $\vec{C}$. First, we have



$$\vec{K} = \frac{\vec{B}+\vec{C}}{2}, \qquad \vec{L} = \frac{\vec{C}+\vec{A}}{2}, \qquad \vec{M} = \frac{\vec{A}+\vec{B}}{2}.$$

Then, we let the circumcenter be the origin. Since the circumcenter is the midpoint of $\overline{A'A}$, $\overline{B'B}$, and $\overline{C'C}$, we have $\vec{A'} = -\vec{A}$, $\vec{B'} = -\vec{B}$, and $\vec{C'} = -\vec{C}$.

We seek a point P such that P is on $\overleftrightarrow{A'K}$, $\overleftrightarrow{B'L}$, and $\overleftrightarrow{C'M}$. From the symmetry of the problem, we expect that we can express $\vec{P}$ as some symmetric expression in terms of $\vec{A}$, $\vec{B}$, and $\vec{C}$. Of course, the simplest outcome would be if P were the origin, but our diagram makes it clear that the lines $\overleftrightarrow{A'K}$, $\overleftrightarrow{B'L}$, and $\overleftrightarrow{C'M}$ don't meet at the center of the circle.

So, we try the next simplest reasonable guess, the orthocenter. In fact, this might have been our first guess if we had taken a close look at the diagram! If $\vec{P} = \vec{A} + \vec{B} + \vec{C}$, we have

$$\vec{A'P} = \vec{P} - \vec{A'} = \vec{A} + \vec{B} + \vec{C} - (-\vec{A}) = 2\vec{A} + \vec{B} + \vec{C},$$
$$\vec{KP} = \vec{P} - \vec{K} = \vec{A} + \vec{B} + \vec{C} - \left(\frac{\vec{B}+\vec{C}}{2}\right) = \frac{2\vec{A}+\vec{B}+\vec{C}}{2}.$$

Success! We have $\vec{A'P} = 2\vec{KP}$, so the orthocenter is on $\overleftrightarrow{A'K}$ for any choice of A, B, and C. Similarly, we can show that the orthocenter is on $\overleftrightarrow{B'L}$ and $\overleftrightarrow{C'M}$, so the three lines are concurrent at the orthocenter.

Once we made the observation that we expect $\vec{P}$ to have a symmetric expression in terms of $\vec{A}$, $\vec{B}$, and $\vec{C}$, we also could have found $\vec{P}$ without further guessing. We start by noting that $\vec{A'K} = \frac{\vec{B}+\vec{C}}{2} - (-\vec{A}) = \vec{A} + \frac{\vec{B}+\vec{C}}{2}$. So, we seek the vector $\vec{P} = t(\vec{A} + \vec{B} + \vec{C})$ such that $\vec{A'P}$ is a scalar multiple of $\vec{A} + \frac{\vec{B}+\vec{C}}{2}$. We have

$$\vec{A'P} = t(\vec{A} + \vec{B} + \vec{C}) - (-\vec{A}) = (t+1)\vec{A} + t\vec{B} + t\vec{C}.$$

This is a scalar multiple of $\vec{A} + \frac{\vec{B}+\vec{C}}{2}$ if and only if $t + 1 = 2t$, which gives us $t = 1$. By essentially the same argument, the point P such that $\vec{P} = \vec{A} + \vec{B} + \vec{C}$ is on the other two lines as well. □

Let's see if we can take some of our tools for triangles up one dimension, and apply them to the three-dimensional analogue of a triangle: a tetrahedron.

Problem 13.12: Each segment from a vertex of a tetrahedron to the centroid of the opposite face of the tetrahedron is called a **median** of the tetrahedron. Prove that the medians of any tetrahedron are concurrent at a point called the **centroid** of the tetrahedron. The centroid of a triangle cuts each median in a 2 : 1 ratio. Can we make a similar statement about the centroid of a tetrahedron?

Solution for Problem 13.12: The vector to the midpoint of a segment is the average of the vectors to the endpoints of the segment. The vector to the centroid of a triangle is the average of the vectors to vertices of the triangle. A reasonable guess is that if there is a point G on all four medians of some tetrahedron $ABCD$, then

$$\vec{G} = \frac{\vec{A}+\vec{B}+\vec{C}+\vec{D}}{4}.$$

Let V be the centroid of face BCD, so

$$\vec{V} = \frac{\vec{B}+\vec{C}+\vec{D}}{3}.$$

To show that G is on $\overline{AV}$, we find $\vec{AG}$ and $\vec{AV}$:

$$\vec{AG} = \vec{G} - \vec{A} = \frac{\vec{A}+\vec{B}+\vec{C}+\vec{D}}{4} - \vec{A} = \frac{\vec{B}+\vec{C}+\vec{D}-3\vec{A}}{4},$$
$$\vec{AV} = \vec{V} - \vec{A} = \frac{\vec{B}+\vec{C}+\vec{D}}{3} - \vec{A} = \frac{\vec{B}+\vec{C}+\vec{D}-3\vec{A}}{3}.$$

So, we have $\vec{AG} = \frac{3}{4}\vec{AV}$, which means that G is the point on $\overline{AV}$ such that $AG/AV = 3/4$. We can similarly show that this choice of G is on each of the other medians, and divides each median in this same ratio. Therefore, just as the medians of a triangle are concurrent at a point that divides each median in a 2 : 1 ratio, the medians of a tetrahedron are concurrent at a point that divides each median in a 3 : 1 ratio. □

Exercises

13.2.1 Find the centroid of the triangle with vertices $(1,4,-1)$, $(-4,5,2)$, and $(6,-1,7)$.

13.2.2 Let $a = BC$, $b = AC$, and $c = AB$ in $\triangle ABC$.

(a) Let D be on $\overline{BC}$ such that $\overline{AD}$ bisects $\angle BAC$. Find $\vec{D}$ in terms of a, b, c, $\vec{A}$, $\vec{B}$, and $\vec{C}$. **Hints:** 104

(b) Let I be the incenter of $\triangle ABC$. Find $\vec{I}$ in terms of a, b, c, $\vec{A}$, $\vec{B}$, and $\vec{C}$.

13.2.3 Show that for any point P and any triangle ABC, the sum of the squares of the sides of the triangle is no greater than 3 times the sum of the squares of the distances from P to each of the vertices of $\triangle ABC$. At what point do we have equality?

13.2.4 Show that if H and O are the orthocenter and circumcenter, respectively, of $\triangle ABC$, then $\vec{H} = \vec{A}+\vec{B}+\vec{C}-2\vec{O}$.

13.2.5 The centroids of the faces of a tetrahedron are $(17,0,3)$, $(-2,13,3)$, $(1,2,15)$, and $(1,-1,2)$. What are its vertices?

13.2.6⋆ Points U, V, and W are the reflections of the circumcenter of $\triangle ABC$ over $\overleftrightarrow{BC}$, $\overleftrightarrow{AC}$, and $\overleftrightarrow{AB}$. Show that $\overleftrightarrow{AU}$, $\overleftrightarrow{BV}$, and $\overleftrightarrow{CW}$ are concurrent. At what special point of $\triangle ABC$ are they concurrent?

13.3⋆ Vectors, Complex Numbers, and Challenging Problems

Our work so far in this chapter probably reminds you of some of the work we did in Chapter 8 with complex numbers. For example, if M is the midpoint of $\overline{AB}$, then $\vec{M} = (\vec{A} + \vec{B})/2$ in vectors and $m = (a + b)/2$ in complex numbers. Also, if $OACB$ is a parallelogram and O is the origin, then $\vec{A} + \vec{B} = \vec{C}$ in vectors and $a + b = c$ in complex numbers. If G is the centroid of $\triangle ABC$, we have $\vec{G} = (\vec{A} + \vec{B} + \vec{C})/3$ and $g = (a + b + c)/3$. And so on. What's going on here?

Let's look back at what a vector in two dimensions is. Vectors in two dimensions are ordered pairs of numbers that have the following operations:

- **Vector addition.** We add vectors by adding the corresponding components:
$$\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a + c \\ b + d \end{pmatrix}.$$
- **Multiplication of a vector by a constant.** We multiply a vector by a constant by multiplying each component of the vector by the constant:
$$k\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ka \\ kb \end{pmatrix}.$$

A complex number is an ordered pair of real numbers, the real part and the imaginary part of the complex number. Complex numbers have these same two operations:

- **Complex number addition.** We have $(a + bi) + (c + di) = (a + c) + (b + d)i$.
- **Multiplication of a complex number by a real constant.** If k is real, we have $k(a + bi) = ka + (kb)i$.

We see that *complex numbers are vectors*, so the similarities between the applications of vectors to geometry and the applications of complex numbers to geometry are not a coincidence!

We also defined multiplication of complex numbers as

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

Note that the dot product is not multiplication of vectors in the same sense, since the dot product of two vectors is not a vector, whereas the product of two complex numbers is itself a complex number. We saw in Chapter 8 that complex number multiplication has some very nice applications to geometry, particularly when dealing with rotations.

As we've seen, we usually use complex numbers and vectors for different types of algebraic problems. For example, we wouldn't use vector notation to express the roots of $x^2 + x + 1$, we don't use the concept of "conjugate" with general vectors in two dimensions, and we don't generally use anything quite like matrices when dealing with complex numbers. Some tools overlap, such as magnitude, which is essentially the same in both.

With most problems that involve complex numbers or more general vectors, it's clear which tool we should use. In this section, we look at a series of very challenging problems in an area where the applications of complex numbers and vectors overlap a great deal: geometry. Some of these problems are best tackled with the machinery we've developed for complex numbers, others with the tools we've used with vectors. Still other problems can be tackled with either, and others with concepts drawn from both.

Problems

Problem 13.13: Consider the points $A(0,12)$, $B(10,9)$, $C(8,0)$, and $D(-4,7)$. There is a unique square S such that each of the four points is on a different side of S. Find the area of S. *(Source: AIME)* **Hints:** 106

Problem 13.14: Points P and Q are on sides $\overline{AB}$ and $\overline{AC}$ of equilateral triangle ABC such that $\overline{PQ} \parallel \overline{BC}$. Show that if M is the midpoint of $\overline{PC}$ and T is the centroid of $\triangle APQ$, then $\triangle TMB$ is a 30-60-90 triangle.

Problem 13.15: Way back on page 185, we helped Petr figure how long his trip to Honolulu would be by finding the distance along the surface of the earth from his home city of Petropavlovsk-Kamchatsky, which has coordinates 52.92° N, 158.49° E, to Honolulu, which has coordinates 21.31° N, 157.83° W. Before he left on his trip, he tried to look up the distance between the two cities on the internet. He found a formula that said the distance between points P and Q on the Earth is αR, where R is the radius of the Earth and α is the angle such that $0 \le \alpha \le \pi$ and

$$\cos\alpha = \cos\lambda_P \cos\lambda_Q \cos(\Delta\beta) + \sin\lambda_P \sin\lambda_Q.$$

Unfortunately, he couldn't find any information about what the symbols λ_P, λ_Q, and $\Delta\beta$ mean. Help Petr by figuring out what this formula is all about. **Hints:** 4

Problem 13.16: Let O be the center of the circumcircle Γ of a regular polygon $A_1A_2A_3\cdots A_{25}A_{26}$ with 26 sides. Let O_1 and O_2 be the reflections of O over $\overline{A_{25}A_1}$ and $\overline{A_2A_6}$, respectively. Prove that $\overline{O_1O_2}$ is equal in length to a side of an equilateral triangle inscribed in Γ. **Hints:** 191

Problem 13.17: Let $ABCDEF$ be a convex hexagon with $AB = BC$, $CD = DE$, and $EF = FA$. Show that the lines containing the altitudes of triangles BCD, DEF, and FAB from C, E, and A, respectively, are concurrent. *(Source: Poland)* **Hints:** 275

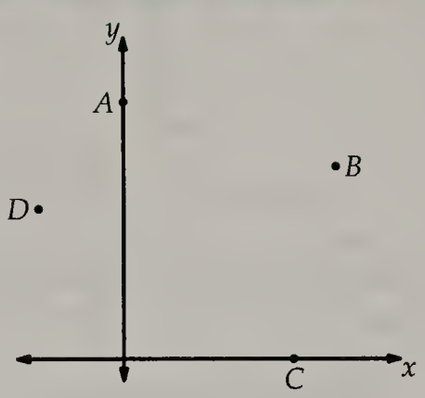
Problem 13.18: Use complex numbers to prove that the area of a triangle with side lengths a, b, and c is

$$\sqrt{s(s-a)(s-b)(s-c)},$$

where $s = (a+b+c)/2$ is the semiperimeter of the triangle. **Hints:** 285

Problem 13.13: Consider the points $A(0,12)$, $B(10,9)$, $C(8,0)$, and $D(-4,7)$. There is a unique square S such that each of the four points is on a different side of S. Find the area of S. *(Source: AIME)*

Solution for Problem 13.13: First, we note that because $\overline{AC}$ and $\overline{BD}$ intersect, each segment connects points on opposite sides of the square. Next, we wonder what additional piece of information we need in order find the side length of the square.



If we knew a direction of the side of the square through A, that would be enough for us to build the whole square. The opposite side has the same direction, so we could use this direction to draw lines through A and C containing the sides of the square. Then, we draw lines through B and D perpendicular to the line through A, and the four intersection points of our four lines give us the vertices of the square.

All this talk about "direction" has us thinking about vectors. Let's see how we might apply vectors to the

problem. We let $\mathbf{u} = \begin{pmatrix} p \\ q \end{pmatrix}$ be a unit vector in a direction of the side of the square through A. Any vector orthogonal to $\begin{pmatrix} p \\ q \end{pmatrix}$ is perpendicular to the side through A, and is therefore in the direction of the side through B. This means that the vector $\mathbf{v} = \begin{pmatrix} -q \\ p \end{pmatrix}$ is in the direction of the side through B. That uses the fact that consecutive sides of a square are perpendicular. Now, we look for a way to use the fact that the sides of a square are the same length.

We can find a vector whose length equals the side length of the square by projecting $\vec{BD}$ onto $\begin{pmatrix} p \\ q \end{pmatrix}$, the direction of the side through A. Similarly, the magnitude of the projection of $\vec{AC}$ onto $\begin{pmatrix} -q \\ p \end{pmatrix}$ must also be the length of a side of the square.

Since $\vec{AC} = \begin{pmatrix} 8-0 \\ 0-12 \end{pmatrix} = \begin{pmatrix} 8 \\ -12 \end{pmatrix}$ and $\vec{BD} = \begin{pmatrix} -4-10 \\ 7-9 \end{pmatrix} = \begin{pmatrix} -14 \\ -2 \end{pmatrix}$, and $\mathbf{u}$ and $\mathbf{v}$ are unit vectors, the projections of $\vec{AC}$ onto $\mathbf{v}$ and of $\vec{BD}$ onto $\mathbf{u}$ are

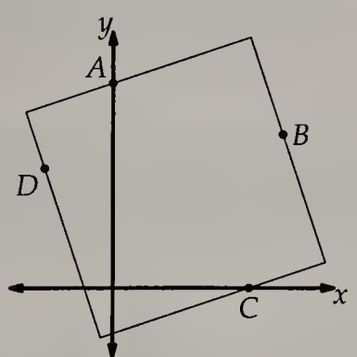
$$\text{proj}_{\mathbf{v}}\vec{AC} = \frac{\vec{AC}\cdot\mathbf{v}}{\|\mathbf{v}\|^2}\mathbf{v} = \frac{-8q-12p}{1^2}\mathbf{v} = (-8q-12p)\mathbf{v},$$

$$\text{proj}_{\mathbf{u}}\vec{BD} = \frac{\vec{BD}\cdot\mathbf{u}}{\|\mathbf{u}\|^2}\mathbf{u} = \frac{-14p-2q}{1^2}\mathbf{u} = (-14p-2q)\mathbf{u}.$$

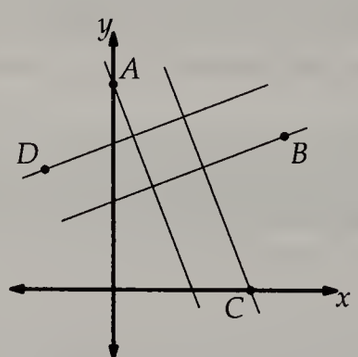
As discussed above, the magnitudes of these projections are equal. We have $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$, so the magnitudes of these projections are $|-8q-12p|$ and $|-14p-2q|$. If these are equal, then we have either $-8q-12p = -14p-2q$ or $-(-8q-12p) = -14p-2q$. The first gives $p = 3q$ and the second gives $13p = -5q$. So, we have

$$\begin{pmatrix} p \\ q \end{pmatrix} = \frac{1}{\sqrt{1^2+3^2}}\begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} p \\ q \end{pmatrix} = \frac{1}{\sqrt{(-5)^2+13^2}}\begin{pmatrix} -5 \\ 13 \end{pmatrix}.$$

(Remember, $\begin{pmatrix} p \\ q \end{pmatrix}$ is a unit vector.) A quick sketch of both cases tells us which we want, and exhibits the interesting result that if there's a second square such that each of the given points is on a different *extended* side of the square.



$$\begin{pmatrix} p \\ q \end{pmatrix} = \frac{1}{\sqrt{1^2+3^2}}\begin{pmatrix} 3 \\ 1 \end{pmatrix}$$



$$\begin{pmatrix} p \\ q \end{pmatrix} = \frac{1}{\sqrt{(-5)^2+13^2}}\begin{pmatrix} -5 \\ 13 \end{pmatrix}$$

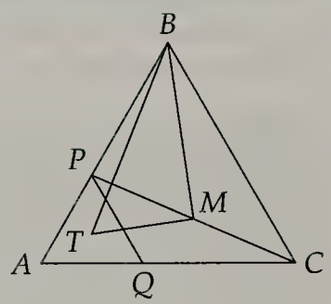
We have $p = 3/\sqrt{10}$ and $q = 1/\sqrt{10}$ in the case in which A, B, C, and D are on the sides of the square themselves (rather than the extensions of the side). The side length of the square equals the magnitude of the projection found above, or $|-8q-12p| = 44/\sqrt{10}$. Therefore, the area of the square is $(44/\sqrt{10})^2 = 193.6$.

(Note that we can also take $\begin{pmatrix} p \\ q \end{pmatrix}$ to be the negatives of the direction vectors found above, and doing so leads to the same two squares.) □

Sidenote: The triangle relationships we developed in the previous section using vectors also hold in complex numbers. Let a, b, and c be complex numbers that correspond to the respective vertices of $\triangle ABC$ in the complex plane. Then, the centroid of $\triangle ABC$ is $(a+b+c)/3$. If the origin is the circumcenter of $\triangle ABC$, then the orthocenter of $\triangle ABC$ is $a+b+c$ and the center of the nine-point circle of $\triangle ABC$ is $(a+b+c)/2$.

Problem 13.14: Points P and Q are on sides $\overline{AB}$ and $\overline{AC}$ of equilateral triangle ABC such that $\overline{PQ} \parallel \overline{BC}$. Show that if M is the midpoint of $\overline{PC}$ and T is the centroid of $\triangle APQ$, then $\triangle TMB$ is a 30-60-90 triangle.

Solution for Problem 13.14: We have complex number tools for dealing with equilateral triangles, so we'll start with complex numbers. As usual, we denote the complex number corresponding to each point with the lowercase version of the uppercase letter used to label the point. We let A be the origin, so equilateral triangle ABC gives us $b = \zeta c$, where ζ is a primitive sixth root of unity. Since $\overline{PQ} \parallel \overline{BC}$, we have $\angle APQ = \angle ABC$ and $\angle AQP = \angle ACB$, which means $\triangle APQ$ is equilateral, too. Therefore, the same rotation that maps C to B also maps Q to P, which means that $p = \zeta q$. Moreover, Q is a dilation of C about the origin, so there's some real constant k such that $q = kc$. Therefore, we have $p = \zeta kc$, and we now have b, p, and q in terms of c.



The only points we haven't tackled yet are M and T. Since M is the midpoint of $\overline{PC}$, we have

$$m = \frac{p+c}{2} = \frac{(\zeta k+1)c}{2}.$$

Since T is the centroid of $\triangle APQ$, we have

$$t = \frac{a+p+q}{3} = \frac{0+kc+\zeta kc}{3} = \frac{(1+\zeta)kc}{3}.$$

Looking at our diagram, it appears that if $\triangle TMB$ is a 30-60-90 triangle, then $\angle TMB = 90°$ and $\angle MTB = 60°$. We could use our usual complex number tools to show separately that $|b-t| = 2|m-t|$ and $\overline{BM} \perp \overline{TM}$, and thereby deduce that $\triangle TMB$ is a 30-60-90 triangle. Instead, we note that we can prove that $\triangle TMB$ is a 30-60-90 triangle in one step if we show that $b - t = 2\zeta(m-t)$, since this equation would tell us that $BT = 2MT$ and that a 60° rotation about T (with the same orientation as the rotation we used to map C to B) maps $\overrightarrow{TM}$ to $\overrightarrow{TB}$. From this, we can deduce that $\triangle TMB$ is a 30-60-90 triangle with $\angle BTM = 60°$ and $\angle BMT = 90°$.

Now that we have a plan, it's time to execute it:

$$2\zeta(m-t) = 2\zeta\left(\frac{(\zeta k+1)c}{2} - \frac{(1+\zeta)kc}{3}\right) = 2c\zeta\left(\frac{3(\zeta k+1)-2(1+\zeta)k}{6}\right) = \frac{c\zeta}{3}(\zeta k+3-2k)$$

and

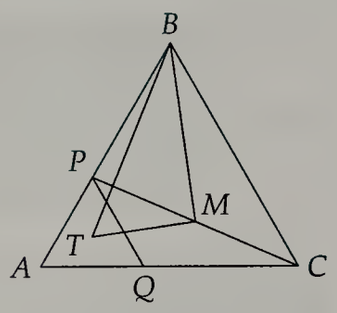
$$b - t = \zeta c - \frac{(1+\zeta)kc}{3} = \frac{3\zeta c - kc - \zeta kc}{3} = \frac{c}{3}(3\zeta - k - \zeta k).$$

To see that these are equal, we'll have to use the fact that $\zeta^2 = \zeta - 1$ for either primitive sixth root of unity:

$$2\zeta(m-t) = \frac{c\zeta}{3}(\zeta k+3-2k) = \frac{c}{3}\left(\zeta^2 k + 3\zeta - 2k\zeta\right) = \frac{c}{3}((\zeta-1)k + 3\zeta - 2k\zeta) = \frac{c}{3}(3\zeta - k - k\zeta).$$

This equals our expression for $b - t$ above, so $b - t = 2\zeta(m-t)$, which means that $\angle BTM = 60°$ and $BT = 2MT$. Therefore, triangle TMB is a 30-60-90 triangle. □

Vectors give us another way to think about some applications of complex numbers to geometry problems. For some people, it is much more natural to think of rotating $\vec{TM}$ by 60° about T to get a vector in the direction of $\vec{BT}$ than it is to think of rotating m by 60° about t to get a complex number that is on the ray from t through b. One way to tie these two ideas together is to associate each complex number with a vector from the origin to the corresponding point in the complex plane. So, $\vec{TM}$ is the vector from t to m in the complex plane. Since $\vec{TM} = \vec{M} - \vec{T}$, we associate $\vec{TM}$ with the complex number $m - t$.



With this in mind, you might realize that you want to show that $b - t = 2\zeta(m - t)$ by thinking, "Triangle TMB is a 30-60-90 triangle if $\angle MTB = 60°$ and $BT = 2MT$. So, if rotating $m - t$ by 60° and then scaling the result by a factor of 2 produces $b - t$, then $\triangle TMB$ is a 30-60-90 triangle. Therefore, we wish to show that $b - t = 2\zeta(m - t)$, where the factor of ζ provides the rotation and the 2 provides the scaling."

Problem 13.15: Way back on page 185, we helped Petr figure how long his trip to Honolulu would be by finding the distance along the surface of the earth from his home city of Petropavlovsk-Kamchatsky, which has coordinates 52.92° N, 158.49° E, to Honolulu, which has coordinates 21.31° N, 157.83° W. Before he left on his trip, he tried to look up the distance between the two cities on the internet. He found a formula that said the distance between points P and Q on the Earth is αR, where R is the radius of the Earth and α is the angle such that $0 \le \alpha \le \pi$ and

$$\cos\alpha = \cos\lambda_P \cos\lambda_Q \cos(\Delta\beta) + \sin\lambda_P \sin\lambda_Q.$$

Unfortunately, he couldn't find any information about what the symbols λ_P, λ_Q, and $\Delta\beta$ mean. Help Petr by figuring out what this formula is all about.

Solution for Problem 13.15: The key to our solution to the earlier problem was finding the angle formed by connecting each city to the center of the Earth. Vectors give us a nice tool for finding angles: the dot product. In order to use the dot product, we'll have to use the rectangular coordinates of the cities. However, the latitude and longitude are in terms of coordinates that are essentially spherical. So, we'll start with spherical coordinates for the cities and convert them into rectangular coordinates. Let the spherical coordinates of cities P and Q be (R, θ_P, ϕ_P) and (R, θ_Q, ϕ_Q), where R is the radius of the Earth. Converting these to rectangular coordinates, city P is at $(R\cos\theta_P \sin\phi_P, R\sin\theta_P \sin\phi_P, R\cos\phi_P)$ and city Q is at $(R\cos\theta_Q \sin\phi_Q, R\sin\theta_Q \sin\phi_Q, R\cos\phi_Q)$. Letting the center of the Earth be the origin, we have

$$\vec{P}\cdot\vec{Q} = \|\vec{P}\|\|\vec{Q}\|\cos\alpha = R^2\cos\alpha,$$

where α is the angle between $\vec{P}$ and $\vec{Q}$. Computing the dot product gives

$$\begin{aligned}
\vec{P}\cdot\vec{Q} &= (R\cos\theta_P \sin\phi_P)(R\cos\theta_Q \sin\phi_Q) + (R\sin\theta_P \sin\phi_P)(R\sin\theta_Q \sin\phi_Q) + (R\cos\phi_P)(R\cos\phi_Q) \\
&= R^2(\cos\theta_P \sin\phi_P \cos\theta_Q \sin\phi_Q + \sin\theta_P \sin\phi_P \sin\theta_Q \sin\phi_Q + \cos\phi_P \cos\phi_Q) \\
&= R^2((\sin\phi_P \sin\phi_Q)(\cos\theta_P \cos\theta_Q + \sin\theta_P \sin\theta_Q) + \cos\phi_P \cos\phi_Q) \\
&= R^2((\sin\phi_P \sin\phi_Q)\cos(\theta_P - \theta_Q) + \cos\phi_P \cos\phi_Q).
\end{aligned}$$

Substituting this into the equation above for $\vec{P}\cdot\vec{Q}$, we have

$$R^2((\sin\phi_P \sin\phi_Q)\cos(\theta_P - \theta_Q) + \cos\phi_P \cos\phi_Q) = R^2\cos\alpha.$$

Dividing by R^2 gives

$$\cos\alpha = \sin\phi_P \sin\phi_Q \cos(\theta_P - \theta_Q) + \cos\phi_P \cos\phi_Q. \tag{13.1}$$

This looks very close to the formula Petr found, which is

$$\cos\alpha = \cos\lambda_P \cos\lambda_Q \cos(\Delta\beta) + \sin\lambda_P \sin\lambda_Q.$$

Right away, we suspect that $\Delta\beta = \theta_P - \theta_Q$. Since the θ coordinates in spherical coordinates are essentially the same as longitude, we can take the $\Delta\beta$ to be the difference in longitude between P and Q.

Turning to the relationship between the ϕ angles and the λ angles, we see that for these angles, Petr's formula has cosines where ours has sines, and vice versa. We unravel this mystery by recalling the relationship between latitude and the ϕ-coordinate of spherical coordinates. If P is in the Southern hemisphere and O is the center of the Earth, then the angle $\overline{OP}$ makes with the upper half of the Earth's axis (ϕ_P in spherical coordinates) is 90° plus the latitude of P. If P is in the Northern hemisphere, then the angle $\overline{OP}$ makes with the upper half of the Earth's axis is 90° minus the latitude of P. Taking λ_P to be the latitude of P, where we let $\lambda_P > 0$ for points in the Northern hemisphere and $\lambda_P < 0$ for points in the Southern hemisphere, we have $\phi_P = 90° - \lambda_P$. Therefore, we have $\cos\phi_P = \sin\lambda_P$ and $\sin\phi_P = \cos\lambda_P$, so Equation (13.1) becomes

$$\cos\alpha = \cos\lambda_P\cos\lambda_Q\cos(\Delta\beta) + \sin\lambda_P\sin\lambda_Q,$$

and the mystery of Petr's formula is solved. □

Problem 13.16: Let O be the center of the circumcircle Γ of a regular polygon $A_1A_2A_3\cdots A_{25}A_{26}$ with 26 sides. Let O_1 and O_2 be the reflections of O over $\overline{A_{25}A_1}$ and $\overline{A_2A_6}$, respectively. Prove that $\overline{O_1O_2}$ is equal in length to a side of an equilateral triangle inscribed in Γ.

Solution for Problem 13.16: Our main clue to start with complex numbers is the regular polygon. We place the problem in the complex plane by letting O be the origin and letting A_k correspond to $e^{k\pi i/13}$ for $1 \le k \le 26$. So, if we let $\omega = e^{\pi i/13}$, then A_k corresponds to ω^k. Since O_1 is the reflection of the origin over $\overline{A_{25}A_1}$, the midpoint of $\overline{A_{25}A_1}$ is the midpoint of $\overline{OO_1}$. Therefore, O_1 is the dilation of this midpoint about the the origin with scale factor 2. The midpoint of $\overline{A_{25}A_1}$ is $(\omega^{25} + \omega)/2$, so O_1 corresponds to twice this number, or $\omega^{25} + \omega$. Similarly, O_2 corresponds to $\omega^6 + \omega^2$.

We wish to show that O_1O_2 equals the length of a side of an equilateral triangle inscribed in the circle. The radius of the circle is 1, so the side length of an equilateral triangle inscribed in the circle is $\sqrt{3}$. We have $O_1O_2 = |\omega^6 + \omega^2 - (\omega^{25} + \omega)|$, so we now must show that

$$|\omega^6 + \omega^2 - \omega^{25} - \omega| = \sqrt{3}.$$

We'll look at $O_1O_2^2$, since we can use the fact that $|z|^2 = z\overline{z}$ to work with our scary expression for O_1O_2. We have

$$O_1O_2^2 = (\omega^6 + \omega^2 - \omega^{25} - \omega)\overline{(\omega^6 + \omega^2 - \omega^{25} - \omega)} = (\omega^6 + \omega^2 - \omega^{25} - \omega)(\overline{\omega}^6 + \overline{\omega}^2 - \overline{\omega}^{25} - \overline{\omega}).$$

That looks plenty terrifying to expand, but before we start doing so, we make a few simplifying observations. First, because $|\omega^k| = 1$, we have $\omega^k\overline{\omega}^k = 1$, which means $\overline{\omega}^k = \frac{1}{\omega^k}$. Since $\omega^{26} = 1$, we have

$$\overline{\omega}^k = \frac{1}{\omega^k} = \frac{\omega^{26}}{\omega^k} = \omega^{26-k}.$$

(See if you can find a geometric explanation for this relationship.) Applying this to the terms in the second factor of our expression for $O_1O_2^2$ gets rid of the conjugates and leaves

$$O_1O_2^2 = (\omega^6 + \omega^2 - \omega^{25} - \omega)(\omega^{20} + \omega^{24} - \omega - \omega^{25}).$$

Next, we note that every vertex of the polygon is diametrically opposite another vertex of the polygon, and is 13 vertices away from its opposite. So, rotating one vertex by π around the origin gives a vertex that is 13 vertices away from the original vertex. In complex numbers, we write this as $\omega^k = -\omega^{k-13}$. Another way to see this is to note that $\omega^{13} = (e^{\pi i/13})^{13} = e^{\pi i} = -1$, so $\omega^k = \omega^{k-13}\cdot\omega^{13} = -\omega^{k-13}$. Applying this relationship to every term in the expression with exponent 13 or greater, we have

$$O_1O_2^2 = (\omega^6 + \omega^2 + \omega^{12} - \omega)(-\omega^7 - \omega^{11} - \omega + \omega^{12}).$$

Now, we're ready to expand, keeping in mind the simplification $\omega^k = -\omega^{k-13}$ whenever we see a power of ω greater than 12. We find

$$\begin{aligned}
O_1O_2^2 &= (\omega^6 + \omega^2 + \omega^{12} - \omega)(-\omega^7 - \omega^{11} - \omega + \omega^{12}) \\
&= \omega^6(-\omega^7 - \omega^{11} - \omega + \omega^{12}) + \omega^2(-\omega^7 - \omega^{11} - \omega + \omega^{12}) \\
&\qquad + \omega^{12}(-\omega^7 - \omega^{11} - \omega + \omega^{12}) - \omega(-\omega^7 - \omega^{11} - \omega + \omega^{12}) \\
&= (-\omega^{13} - \omega^{17} - \omega^7 + \omega^{18}) + (-\omega^9 - \omega^{13} - \omega^3 + \omega^{14}) \\
&\qquad + (-\omega^{19} - \omega^{23} - \omega^{13} + \omega^{24}) + (\omega^8 + \omega^{12} + \omega^2 - \omega^{13}) \\
&= (1 + \omega^4 - \omega^7 - \omega^5) + (-\omega^9 + 1 - \omega^3 - \omega) \\
&\qquad + (\omega^6 + \omega^{10} + 1 - \omega^{11}) + (\omega^8 + \omega^{12} + \omega^2 + 1) \\
&= 4 + \omega^{12} - \omega^{11} + \omega^{10} - \omega^9 + \omega^8 - \omega^7 + \omega^6 - \omega^5 + \omega^4 - \omega^3 + \omega^2 - \omega.
\end{aligned}$$

That's quite a compelling expression! We can simplify it either by recognizing it as geometric series with common ratio $-\omega$, or by recalling the factorization

$$\omega^{13} + 1 = (\omega + 1)(\omega^{12} - \omega^{11} + \omega^{10} - \omega^9 + \omega^8 - \omega^7 + \omega^6 - \omega^5 + \omega^4 - \omega^3 + \omega^2 - \omega + 1).$$

Since we have $\omega^{13} = -1$ but $\omega \neq 1$, the factorization above gives us

$$0 = \omega^{12} - \omega^{11} + \omega^{10} - \omega^9 + \omega^8 - \omega^7 + \omega^6 - \omega^5 + \omega^4 - \omega^3 + \omega^2 - \omega + 1,$$

so $\omega^{12} - \omega^{11} + \omega^{10} - \omega^9 + \omega^8 - \omega^7 + \omega^6 - \omega^5 + \omega^4 - \omega^3 + \omega^2 - \omega = -1$. Substituting this into our expression for $O_1O_2^2$ gives $O_1O_2^2 = 4 - 1 = 3$, so $O_1O_2 = \sqrt{3}$, as desired. □

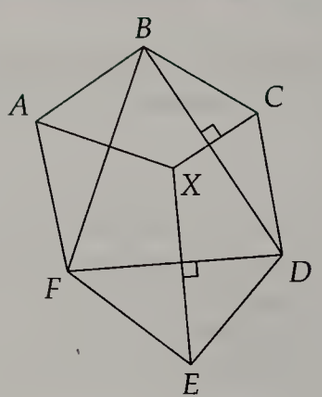
Concept: Sometimes persistent algebraic manipulation can tackle hard geometry problems when your geometric instincts fail you.

That said, don't always reach for complex numbers or vectors right away when you see a geometry problem.

WARNING!! Unless you very quickly see that vectors or complex numbers can be used to solve a particular geometry problem, you should usually spend some time trying to apply geometric tools to the problem before reaching for algebraic tools.

Problem 13.17: Let $ABCDEF$ be a convex hexagon with $AB = BC$, $CD = DE$, and $EF = FA$. Show that the lines containing the altitudes of triangles BCD, DEF, and FAB from C, E, and A, respectively, are concurrent. *(Source: Poland)*

Solution for Problem 13.17: The condition for perpendicularity in vectors is a little easier to work with than the condition we have for complex numbers, so we'll try applying vectors to the problem. We'll let X be the intersection of the altitudes from C and E, and our goal is to prove that X is on the altitude from A, as well. Since X is on the altitudes from C and E to $\overline{BD}$ and $\overline{DF}$, respectively, we have $\vec{CX} \cdot \vec{BD} = 0$ and $\vec{EX} \cdot \vec{DF} = 0$. We'd like to prove that $\vec{AX} \cdot \vec{FB} = 0$. Since there are many points in common in each of these statements, we write each vector as a vector difference, hoping to be able to find some convenient ways to combine the equations:



$$\begin{aligned}
(\vec{X} - \vec{C}) \cdot (\vec{D} - \vec{B}) &= 0, \\
(\vec{X} - \vec{E}) \cdot (\vec{F} - \vec{D}) &= 0.
\end{aligned}$$

It's tempting at this point to choose X to be the origin, but the conditions $AB = BC$, $CD = DE$, and $EF = FA$ give us pause. It's not clear how we would use these conditions if X were the origin. (That said, it is possible to finish the problem with X as the origin! See if you can find such a solution on your own.) Rather than choose our origin right away, we'll continue with the problem, hoping that we'll find a choice of origin that allows us to use the given equal lengths effectively.

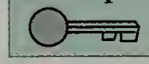
Our hopes of combining the two dot products above to show that $(\vec{X}-\vec{A})\cdot(\vec{B}-\vec{F}) = 0$ appear to be dashed—there are no instances of $\vec{A}$ in the first two equations, and it's not clear how to introduce $\vec{A}$ and get rid of $\vec{E}$ and $\vec{C}$. We do at least see that we'll get some cancellation if we expand and add the left sides, since the $\vec{X}\cdot\vec{D}$ in the first equation will cancel with the $-\vec{X}\cdot\vec{D}$ in the second. But the rest will be a mess.

The symmetry of the problem suggests that we go ahead and include the desired dot product:

$$(\vec{X}-\vec{C})\cdot(\vec{D}-\vec{B}) = 0,$$
$$(\vec{X}-\vec{E})\cdot(\vec{F}-\vec{D}) = 0,$$
$$(\vec{X}-\vec{A})\cdot(\vec{B}-\vec{F}) = k.$$

Our goal is to show that $k = 0$. Again, note that we've set up these equations to have a convenient symmetry. This symmetry would have been obscured if we had written the equations in the equivalent forms:

$$(\vec{C}-\vec{X})\cdot(\vec{D}-\vec{B}) = 0,$$
$$(\vec{X}-\vec{E})\cdot(\vec{D}-\vec{F}) = 0,$$
$$(\vec{A}-\vec{X})\cdot(\vec{F}-\vec{B}) = k.$$



Concept: When you have options for how to set up parts of a problem, do so in a way that allows you to exploit symmetry where possible.

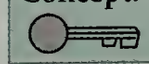
Returning to the equations

$$(\vec{X}-\vec{C})\cdot(\vec{D}-\vec{B}) = 0,$$
$$(\vec{X}-\vec{E})\cdot(\vec{F}-\vec{D}) = 0,$$
$$(\vec{X}-\vec{A})\cdot(\vec{B}-\vec{F}) = k,$$

we see that if we expand the dot products on the left and add all the results, all of the terms with $\vec{X}$ will cancel, and we'll be left with

$$-\vec{C}\cdot(\vec{D}-\vec{B}) - \vec{E}\cdot(\vec{F}-\vec{D}) - \vec{A}\cdot(\vec{B}-\vec{F}) = k.$$

We'd like to show that the left side is 0, but in this form, it's not clear how to. Expanding everything on the left side won't offer any cancellation. However, we could group the terms on the left side by $\vec{B}$, $\vec{D}$, and $\vec{F}$ instead of by $\vec{A}$, $\vec{C}$, and $\vec{E}$.



Concept: A different way of looking at the same information can offer important insights.

Grouping the terms by $\vec{B}$, $\vec{D}$, and $\vec{F}$ gives

$$k = -\vec{C}\cdot(\vec{D}-\vec{B}) - \vec{E}\cdot(\vec{F}-\vec{D}) - \vec{A}\cdot(\vec{B}-\vec{F}) = \vec{B}\cdot(\vec{C}-\vec{A}) + \vec{D}\cdot(\vec{E}-\vec{C}) + \vec{F}\cdot(\vec{A}-\vec{E}). \tag{13.2}$$

OK, maybe that's not so helpful, either. We seem stuck.

Concept: When stuck on a problem, focus information in the problem that you haven't used yet.

We haven't used the side equalities. What can we deduce from $AB = BC$? We can deduce that B is on the perpendicular bisector of $\overline{AC}$. Similarly, D is on the perpendicular bisector of $\overline{CE}$ and F is on the perpendicular bisector of $\overline{AE}$. Aha! There's a point on all three of these perpendicular bisectors—the circumcenter of $\triangle ACE$! Looking back at our expressions above, we see that this observation solves the problem. If we choose the circumcenter of $\triangle ACE$ to be the origin, then $\overleftrightarrow{OB} \perp \overline{AC}$, $\overleftrightarrow{OD} \perp \overline{CE}$, and $\overleftrightarrow{OF} \perp \overline{EA}$, so we have

$$\vec{B} \cdot \vec{AC} = \vec{D} \cdot \vec{CE} = \vec{F} \cdot \vec{EA} = 0.$$

Applying these to Equation (13.2), we have

$$k = \vec{B} \cdot (\vec{C} - \vec{A}) + \vec{D} \cdot (\vec{E} - \vec{C}) + \vec{F} \cdot (\vec{A} - \vec{E}) = \vec{B} \cdot \vec{AC} + \vec{D} \cdot \vec{CE} + \vec{F} \cdot \vec{EA} = 0.$$

Since $\vec{AX} \cdot \vec{FB} = k = 0$, point X is also on the altitude from A to $\overline{FB}$. Therefore, the three altitudes described in the problem are concurrent. □

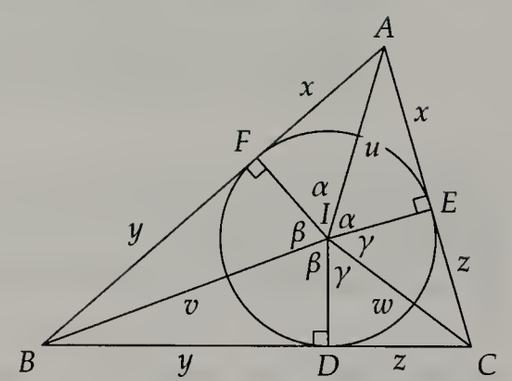
Problem 13.18: Prove that the area of a triangle with side lengths a, b, and c is

$$\sqrt{s(s-a)(s-b)(s-c)},$$

where $s = (a+b+c)/2$ is the semiperimeter of the triangle.

Sidenote: Art of Problem Solving Community member Miles Dillon Edwards produced the following dazzling proof of Heron's formula while still in high school. His proof was published in the *American Mathematical Monthly* in 2008.

Solution for Problem 13.18: We need to prove a formula for the area of a triangle that involves the semiperimeter. We already have one such formula: the area of a triangle equals its semiperimeter times the radius of its incircle (which we call the inradius). So, we start with the diagram at right, in which the angle bisectors of $\triangle ABC$ meet at the incenter, I, as shown. We connect the incenter to the three vertices and to the points at which the incircle is tangent to the sides of the triangle, forming three pairs of congruent right triangles, such as $\triangle AFI$ and $\triangle AEI$. We label the pairs of equal angles at I as shown, as well as the equal tangents from the vertices.



We let the radius of the circle be r, and we label the distances from I to the vertices as shown. Adding all the angles at I gives us $\alpha + \beta + \gamma = \pi$, since all six angles at I must sum to 2π. That's a pretty convenient angle. Multiplying complex numbers leads to adding angles. Maybe we can tackle this problem with complex numbers.

We'll need complex numbers with arguments α, β, and γ. In right triangle IEA, we have $r^2 + x^2 = u^2$ and $\tan\alpha = \frac{x}{r}$, so the complex number $r + xi$ has magnitude u and argument α. In other words, we have $r + xi = ue^{i\alpha}$. Similarly, we have $r + yi = ve^{i\beta}$ from $\triangle BFI$ and $r + zi = we^{i\gamma}$ from $\triangle CDI$.

Our purpose in finding these complex numbers was to multiply them, and thereby add their arguments:

$$(r + xi)(r + yi)(r + zi) = (ue^{i\alpha})(ve^{i\beta})(we^{i\gamma}) = uvwe^{i(\alpha+\beta+\gamma)}.$$

Since $\alpha + \beta + \gamma = \pi$, we have

$$(r + xi)(r + yi)(r + zi) = uvwe^{\pi i} = -uvw,$$

which means that $(r + xi)(r + yi)(r + zi)$ is real. Therefore, the imaginary part of this expression is 0. Expanding this product yields gives

$$\text{Im}\left((r + xi)(r + yi)(r + zi)\right) = r^2(x + y + z) - xyz.$$

Since this equals 0, we have

$$r = \sqrt{\frac{xyz}{x + y + z}}.$$

Since the perimeter of $\triangle ABC$ is $2(x + y + z)$, we have $x + y + z = s$. Letting $a = BC$, $b = AC$, and $c = AB$, we have $a = y + z$, $b = x + z$, and $c = x + y$. Therefore, we have $x = s - y - z = s - (y + z) = s - a$. Similarly, we find $y = s - b$ and $z = s - c$. So, we have

$$r = \sqrt{\frac{(s - a)(s - b)(s - c)}{s}}.$$

We finish by noting that the area of $\triangle ABC$ is the sum of the areas of $\triangle IAB$, $\triangle IAC$, and $\triangle IBC$, which is

$$\frac{ra}{2} + \frac{rb}{2} + \frac{rc}{2} = r\left(\frac{a + b + c}{2}\right) = rs = \sqrt{s(s - a)(s - b)(s - c)}.$$

This formula is known as **Heron's Formula**. □

13.4 Summary

Important: If M is the midpoint of $\overline{AB}$, then $\vec{M} = (\vec{A} + \vec{B})/2$.

Important: The medians of any triangle are concurrent at a point called the **centroid**. If the centroid of $\triangle ABC$ is G, then $\vec{G} = (\vec{A} + \vec{B} + \vec{C})/3$.

Important: The altitudes of any triangle are concurrent at a point called the **orthocenter**. Let H be the orthocenter of $\triangle ABC$. If the circumcenter of $\triangle ABC$ is the origin, then $\vec{H} = \vec{A} + \vec{B} + \vec{C}$.

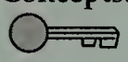
Things To Watch Out For!

WARNING!!

- We have $\vec{AB} = \vec{B} - \vec{A}$, not $\vec{AB} = \vec{A} - \vec{B}$.
- Let H be the orthocenter of $\triangle ABC$. We proved that if the circumcenter of $\triangle ABC$ is the origin, then $\vec{H} = \vec{A} + \vec{B} + \vec{C}$. If the circumcenter of $\triangle ABC$ is not the origin, then we cannot conclude that $\vec{H} = \vec{A} + \vec{B} + \vec{C}$.

Problem Solving Strategies

Concepts:



- Choosing your origin wisely can greatly simplify the application of vectors to geometry problems.
- When stuck on a problem, look back at the problem for any information you haven't used.
- When applying vectors to a problem involving a triangle, choosing the circumcenter as the origin can greatly simplify the problem.
- Don't overlook "guess-and-check" as a strategy, even in advanced problems. Guessing and then confirming that your guess is correct has a long and glorious history in mathematics and science.
- Precise diagrams can help you discover important relationships in geometry problems.
- Sometimes persistent algebraic manipulation can tackle hard geometry problems when your geometric instincts fail you.
- When you have options for how to set up parts of a problem, do so in a way that allows you to exploit symmetry where possible.
- A different way of looking at the same information can offer important insights.

Challenge Problems

13.19 Show that for any points W, X, Y, and Z, we have $WX^2 + YZ^2 - XY^2 - ZW^2 = 2\vec{WY} \cdot \vec{ZX}$.

13.20 Use vectors to show that the sum of the squares of the side lengths of a quadrilateral equals the sum of the squares of the diagonals of the quadrilateral if and only if the quadrilateral is a parallelogram.

13.21 In tetrahedron $ABCD$, points M, N, P, Q, R, and S are the midpoints of $\overline{AB}$, $\overline{CD}$, $\overline{AC}$, $\overline{BD}$, $\overline{AD}$, and $\overline{BC}$, respectively. Show that $\overleftrightarrow{MN}$, $\overleftrightarrow{PQ}$, and $\overleftrightarrow{RS}$ are concurrent. **Hints:** 57

13.22 Let the diagonals of $ABCD$ meet at X. Show that if

$$AB^2 + BC^2 + CD^2 + DA^2 = 2(AX^2 + BX^2 + CX^2 + DX^2).$$

then either the diagonals of $ABCD$ are perpendicular or one of the diagonals bisects the other. **Hints:** 15

13.23 Mr. Assumption was trying to find the orthocenter of the triangle with vertices $(4, -3)$, $(2, -1)$, and $(-1, 5)$. Here's what he wrote: "If $\triangle ABC$ is a triangle, then the orthocenter H satisfies $\vec{H} = \vec{A} + \vec{B} + \vec{C}$. So, we get the orthocenter by simply adding the coordinates: $(4 + 2 - 1, -3 - 1 + 5) = (5, -1)$." What mistake did Mr. Assumption make?

13.24 Let O be the circumcenter of $\triangle ABC$. Let M be the midpoint of $\overline{AB}$, and let X be the centroid of $\triangle ACM$. Show that $\overleftrightarrow{CM} \perp \overleftrightarrow{OX}$ if and only if $AB = AC$.

13.25 Let $ABCD$ be a cyclic quadrilateral (a quadrilateral that is inscribed in a circle). Show that the orthocenters of ABC, BCD, CDA, and DAB are the vertices of a quadrilateral that is congruent to $ABCD$.

13.26 Let O, H, and R be the circumcenter, orthocenter, and circumradius, respectively, of $\triangle ABC$.

(a) Show that $OH^2 = R^2(3 + 2(\cos 2A + \cos 2B + \cos 2C))$. **Hints:** 224

(b) Show that in any $\triangle ABC$, we have $\cos^2 A + \cos^2 B + \cos^2 C \geq \frac{3}{4}$. **Hints:** 82

13.27 Let point X be any point inside $\triangle ABC$. Through the midpoints of $\overline{AB}$, $\overline{BC}$, and $\overline{CA}$, we draw lines that are parallel to $\overline{XC}$, $\overline{XA}$, and $\overline{XB}$, respectively. Show that these three lines are concurrent. **Hints:** 64

13.28 Points U, V, W, X, Y, and Z are the centroids of triangles ABC, BCD, CDE, DEF, EFA, and FAB, respectively, in convex hexagon $ABCDEF$. Show that each pair of opposite sides of hexagon $UVWXYZ$ are parallel and equal in length.

13.29 In tetrahedron $ABCD$, the angle $\angle BDC$ is a right angle and the foot of the perpendicular from D to ABC is the intersection of the altitudes of $\triangle ABC$. Show that $\angle ADB = \angle ADC = 90°$. *(Source: IMO)* **Hints:** 215, 258

13.30 Harrison wants to construct a hexagon such that the midpoints of its edges have coordinates $(0,0)$, $(5,0)$, $(6,2)$, $(4,4)$, $(2,4)$, and $(-2,3)$, in no particular order. Can he do this? If so, how? **Hints:** 232

13.31 Describe the set of points X that satisfy $\vec{AX} \cdot \vec{CX} = \vec{CB} \cdot \vec{AX}$, where A, B, and C are three different points.

13.32 Regular polygon $P_1P_2P_3\cdots P_n$ is inscribed in a circle with center C and radius r. Show that for any point X, we have

$$(P_1X)^2 + (P_2X)^2 + (P_3X)^2 + \cdots + (P_nX)^2 = n(r^2 + CX^2).$$

13.33 Vertex A of the acute triangle ABC is equidistant from the circumcenter O and the orthocenter H of $\triangle ABC$. Determine all possible values for the measure of angle A. *(Source: IMO)* **Hints:** 234

13.34 Three vertices of a regular octahedron are at $(2,6,-8)$, $(6,4,-4)$, and $(4,8,0)$. Find the coordinates of the other three vertices of the octahedron. **Hints:** 114, 43, 173

13.35 Let a and b be complex numbers. Show that the area of the parallelogram in the complex plane with consecutive vertices at the origin, a, and b, is $\left|\frac{1}{2}(a\bar{b} - \bar{a}b)\right|$. **Hints:** 167, 196

13.36★ In tetrahedron $ABCD$, we have $AB = CD = 6$, $AC = BD = 8$, and $AD = BC$. Let X be the centroid of $\triangle BCD$ and Y be the centroid of $\triangle ABC$. Find AD if $\overline{AX} \perp \overline{DY}$. **Hints:** 251, 197

13.37★ A fixed point P is given inside a sphere with center O. Three mutually perpendicular rays (meaning each pair of rays is perpendicular) from P intersect the sphere at points U, V, and W. Point Q is the vertex diagonally opposite to P in the parallelepiped with edges $\overline{PU}$, $\overline{PV}$, and $\overline{PW}$. Show that OQ is the same for all possible choices of the three mutually perpendicular rays from P. *(Source: IMO)*

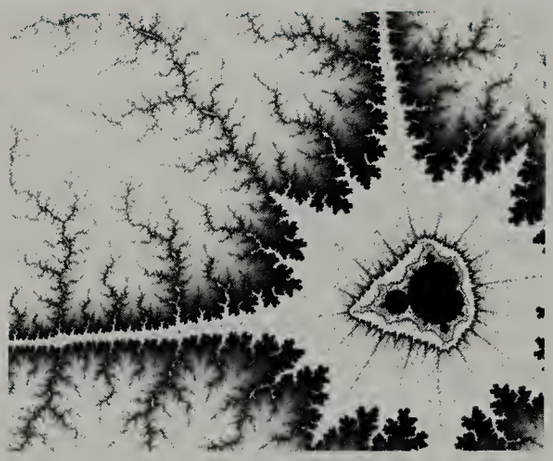
13.38★ $\overline{WY}$ and $\overline{XZ}$ intersect at O. Points G_1 and G_2 are the centroids of triangles WXO and YZO, and points H_1 and H_2 are the orthocenters of XYO and ZWO. Show that $\overline{G_1G_2} \perp \overline{H_1H_2}$. **Hints:** 247, 13, 28, 200

13.39★ The altitudes of tetrahedron $ABCD$ are extended externally beyond A, B, C, and D to points E, F, G, and H, respectively, such that $AE = k/h_a$, $BF = k/h_b$, $CG = k/h_c$, and $HD = k/h_d$, where k is a constant and h_a denotes the length of the altitude of $ABCD$ from vertex A, and likewise for h_b, h_c and h_d. Prove that the centroid of $EFGH$ is also the centroid of $ABCD$.*(Source: Canada)* **Hints:** 100, 140

13.40★ The angles of convex hexagon $ABCDEF$ are equal. Prove that $|AB - DE| = |BC - EF| = |CD - FA|$. **Hints:** 180, 42

13.41★ Let D, E, and F be on $\overline{BC}$, $\overline{AC}$, and $\overline{AB}$, respectively, such that $\overline{AD}$, $\overline{BE}$, and $\overline{CF}$ be the angle bisectors of $\triangle ABC$. If $\angle EDF = 90°$, then find all possible values of $\angle BAC$. *(Source: USAMO)* **Hints:** 201, 76

13.42★ Triangle ABC is isosceles with $AB = AC$ and $BC = 1$. D on $\overline{AB}$ such that $AD = 1$ and $\angle DAC = \pi/7$. Find CD. **Hints:** 154

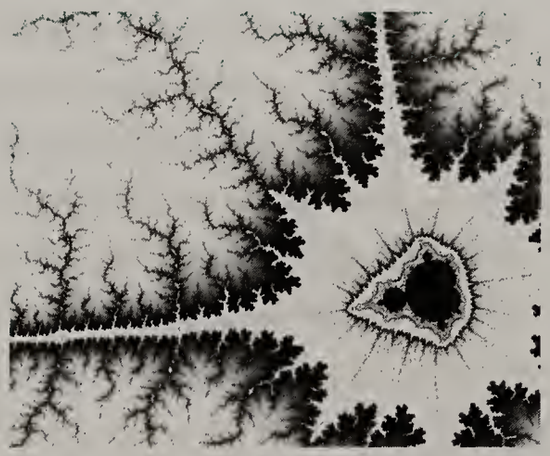


References

In addition to the sources cited below, there are links to many useful websites on the book's links page at

`http://www.artofproblemsolving.com/BookLinks/Precalculus/links.php`

1. T. Andreescu, D. Andrica, *Complex Numbers from A to...Z*, Birkhäuser, 2006.
2. C. Bradley, A. Gardiner, *Plane Euclidean Geometry: Theory and Problems*, The United Kingdom Mathematics Trust, 2005.
3. C. Bradley, *The Algebra of Geometry*, Highperception, 2007.
4. M. Crawford, R. Rusczyk, *Intermediate Algebra*, AoPS Incorporated, 2008.
5. M. Edwards, "A Proof of Heron's Formula," *American Mathematical Monthly*, **114**, no. 10, p. 937.
6. A. Engel, *Problem-Solving Strategies*, Springer-Verlag, 1998.
7. L. Larson, *Problem-Solving Through Problems*, Springer-Verlag, 1983.
8. S. Lehoczky, R. Rusczyk, *the Art of Problem Solving, Volume 2: and Beyond*, 7th edition, AoPS Incorporated, 2006.
9. E. Weisstein "Cycloid." From *MathWorld*—A Wolfram Web Resource. `http://mathworld.wolfram.com/Cycloid.html`
10. P. Zeitz, *The Art and Craft of Problem Solving*, 2nd edition, John Wiley & Sons, Inc., 2007.



Hints to Selected Problems

1. Complete the square. (That means write the quadratic in the form $a(x-h)^2+k$ for some constants a, h, and k.)
2. All five choices are nonreal, so one of the factors of the polynomial is a quadratic with nonreal roots. Which of the given choices can be a root of a monic quadratic with integer coefficients?
3. Yes, you'll need a *lot* of algebra. Two factorizations that will be helpful are $x^2+2xy+y^2=(x+y)^2$ and $x^2-y^2=(x-y)(x+y)$.
4. What vector tool involves cosines?
5. In solving the first part, you should have found equations for p and q in terms of s and t (and possibly in terms of some constants you defined). Prove that you can solve these equations for s and t in terms of p and q.
6. Let $\mathbf{v}=a\mathbf{i}+b\mathbf{j}+c\mathbf{k}$ and $\vec{F}=s\mathbf{i}+t\mathbf{j}+u\mathbf{k}$. Find s, t, and u in terms of a, b, and c.
7. Find other angles in terms of θ. A lot of triangles share side lengths in this problem. What triangle law might you use to take advantage of this in order to create an equation for θ?
8. Draw a diagram for the situation after $\mathcal{B}$ has rotated θ counterclockwise about $\mathcal{A}$. By what angle has the red spot rotated about the center of $\mathcal{B}$?
9. Let $AC=BC=1$. Find other sides in terms of the angles. Keep going until you discover something interesting.
10. If the hole is drilled along the z-axis, then are there any restrictions on θ when describing red points that remain? Are there restrictions on ϕ?
11. How can you get information about squares from the given equation?
12. Each term is a power of $2z$.
13. To deal with h_2, draw a couple altitudes in $\triangle WOZ$ (h_2 is the orthocenter of this triangle), and look for similar triangles.
14. Note that ω is a primitive cube root of unity. So, what quadratic has ω as a root? Use this to find an expression for ω^2.

15. What choice of origin will greatly simplify the right side of the given equation?

16. We learned a tool in this chapter for dealing with $\cos n\theta$ and $\sin n\theta$.

17. Let the center of the square be the origin and let k be the x-axis.

18. A quadrilateral can be cut into two triangles.

19. You (should have) proved similar identities as Exercises. Are any of them helpful here?

20. Let $t = \tan 20°$. Write all three tangents in terms of t.

21. Write $\cos^2 B + \cos^2 C$ in terms of $\cos 2B$ and $\cos 2C$. Don't forget that $B + C = 180° - A$!

22. What does $|z| = 1$ tell us about the location of z in the complex plane?

23. Let θ be the angle that $\overline{OX}$ makes with the positive x-axis. What point has the same x-coordinate as A? What point has the same y-coordinate as A?

24. The angle you found in the first hint isn't the only angle with that tangent! Tangent has a period of π.

25. Note that $\cos 7\theta = 1$. Write $\cos 7\theta$ in terms of $\cos\theta$. Let $x = \cos\theta$. You should have a degree 7 polynomial. This polynomial must have 7 roots. Guess what they are!

26. Let α be the sum of the first two inverse cotangents and let β be the sum of the second two. Now does the problem look more manageable?

27. Let R be the radius of the Earth. Find VC, CO, and CD in terms of R.

28. Note that $\triangle WOZ$ and $\triangle XOY$ have an angle in common. So, relating h_2 to this angle may help you relate h_2 to h_1.

29. Time to take the training wheels off. You're on your own on this one!

30. Use trig identities to express the ratio of tangents in terms of $\sin A$ and $\sin B$.

31. Put the problem on the unit circle.

32. Because $\mathbf{v}$ and $\mathbf{w}$ are in a plane generated by $\mathbf{a}$ and $\mathbf{b}$, both $\mathbf{v}$ and $\mathbf{w}$ are linear combinations of $\mathbf{a}$ and $\mathbf{b}$. Write equations for $\mathbf{v}$ and $\mathbf{w}$ using this fact.

33. Under what conditions does a quadratic have real roots?

34. Let $c = \cos\theta$. Find a quadratic in z that has coefficients in terms of c.

35. Think geometrically. What points satisfy $|z_1 - 5 + 3i| \leq 4$?

36. Letting $\mathbf{e} = \mathbf{a} \times \mathbf{b}$ might help you see the first step better.

37. How can we relate sine to an exponential?

38. How is a real number related to its conjugate?

39. Factor the first two terms and the last two terms.

40. What is the distance between $x + xi$ and i?

41. Suppose we have $\mathbf{M}\mathbf{u}_i = \mathbf{v}_i$ for $i = 1, 2, 3$. If $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are linearly independent, then must $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ be linearly independent?

42. Try complex numbers. Assign variables to the side lengths. Find $b - a$, $c - b$, $d - c$, etc., in terms of the side lengths and ζ, where ζ is a primitive sixth root of unity.

43. Can you find the center of the octahedron?

44. When you need to get rid of sines and cosines, consider the identity $\sin^2\theta + \cos^2\theta = 1$.

45. Let $\mathbf{v}$ be the vector from a point on the line to $(4, -2)$. Onto what vector can we project $\mathbf{v}$ to get a vector whose length is the desired distance?

46. Let O be the origin, and let U, V, and W be points such that $\vec{U} = \mathbf{u}$, $\vec{V} = \mathbf{v}$, and $\vec{W} = \mathbf{w}$. How are $\overline{OV}$ and $\overline{VU}$ related?

47. Two useful tactics with quadratics are factoring and completing the square. Try them both!

48. This is the "polar coordinates" section!!

49. Let $x = AC$ and $\theta = \angle C$. Find other angles and lengths in terms of x and θ. Try to build an equation for x.

50. Can you find t in terms of x and y?

51. Under what conditions is there a vector $\mathbf{w}$ besides $\mathbf{0}$ such that $\mathbf{Mw} = \mathbf{0}$?

52. If $\sin t + i\cos t$ were replaced with something of the form $\cos\theta + i\sin\theta$, you could use de Moivre's Theorem.

53. What graph is produced when the graph of $y = f(x)$ is shifted 6 to the right, and then the result scaled horizontally by a factor of $1/2$?

54. When it comes to matrix multiplication, what matrix plays the role that 1 plays when multiplying real numbers?

55. If $\sin\theta = \sin\phi$, then how are θ and ϕ related on the unit circle! (Note: they don't necessarily have to differ by a multiple of 2π!)

56. Write both sides in terms of complex numbers. There is a sum on the greater side of the Triangle Inequality. (Don't be afraid of a little algebra!)

57. What's a good guess for the point on $\overleftrightarrow{MN}$ at which the lines meet?

58. After writing the sum in a concise form, can you write $P(x)$ as the product of two polynomials whose roots you know?

59. Sometimes area can be a useful problem solving tool in problems that don't even appear to be about area.

60. Note that $z = 0$ if and only if $|z| = 0$.

61. What is $z\bar{z}$ if the center of the circle is the origin and the radius of the circle is 1?

62. Where do the graphs of f and g cross the x-axis?

63. What angle equals $\angle HBX$?

64. Draw a very precise diagram. How does $\vec{CX}$ appear to be related to $\vec{VM}$?

65. What trig identity does the left side resemble?

66. What is $\frac{y}{x+\sqrt{x^2+y^2}}$ in terms of r and θ?

67. Set the desired expression equal to x.

68. Your trig identity toolbox will get a lot of work on this one! For starters, $\sin^2\theta + \cos^2\theta = 1$ isn't the only identity that has $\cos^2\theta$.

69. What do you know about $w\overline{w_1}$? What can you multiply $w_1w_2 + w_2w_3 + w_3w_1$ by to introduce expressions you can evaluate?

70. How is the magnitude of a complex number related to the conjugate of the number?

71. Do you see any cases in which the determinant equals 0? What factors do these cases suggest?

72. Notice that 40° and 80° are 20° from 60°.

73. What is $xf(x)$? Can you use the result to find a simpler expression for $f(x)$? How is that expression related to z?

74. The left side is a sum of angles. When working with complex numbers, when do we end up adding angles?

75. What are the possible values of $\det(\mathbf{A})$?

76. Let A be the origin. Find an expression for $\vec{B} \cdot \vec{C}$.

77. You're probably stuck dealing with $e^{i\alpha}e^{i\beta} + e^{i\alpha}e^{i\gamma} + e^{i\beta}e^{i\gamma}$. If so, what can you multiply the given equation by to get something that looks *almost* like $e^{i\alpha}e^{i\beta} + e^{i\alpha}e^{i\gamma} + e^{i\beta}e^{i\gamma}$? And how is the result of that multiplication related to $e^{i\alpha}e^{i\beta} + e^{i\alpha}e^{i\gamma} + e^{i\beta}e^{i\gamma}$?

78. Build right triangles. Lots of them! What do you know about the acute angles that have the center of the circle as the vertex?

79. Where else in math have you seen the sum of the squares of two real numbers set equal to 1?

80. Let the points be A, B, C, respectively. By what angle must we rotate B about A to get C?

81. What vector is parallel to $B_1\vec{B}_2$?

82. In part (a), we proved an equation involving $\cos 2A$. In this part, we have a statement about $\cos^2 A$. How can we relate the two?

83. Suppose the calculator evaluates sine in radians, and that the professor entered a measure t that was in degrees. What should the professor have entered instead of simply "t"?

84. $\sin^2\theta + \cos^2\theta = 1$

85. If you're stuck with a ratio that you'd like to show is imaginary, here's a hint. If $\frac{z_1}{z_2} = \frac{z_3}{z_4} = k$, then what does $\frac{z_1 b + z_3 c}{z_2 b + z_4 c}$ equal?

86. The series $\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \cdots + \frac{1}{(n-1)\cdot n}$ telescopes, because $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$.

87. If you're stuck dealing with $w\bar{z} + z\bar{w}$, try computing this expression for a few choices of w and z. Notice anything interesting?

88. Let $S = w + 2w^2 + 3w^3 + \cdots + 9w^9$. What is wS?

89. Once you start dealing with the fourth powers, the algebra gets a little hairy. But don't forget what $\omega^2 + \omega + 1$ is when ω is a primitive cube root of unity!

90. Geometry: Start with a regular pentagon. Let the side length of the pentagon be $2x$. Add diagonals and drop altitudes to build some 18-72-90 right triangles.

91. Find the possible values of p first. Then, for each possible value, think about where the graph crosses the x-axis.

92. No luck with the first hint? Let's try another approach. Let $x = BD$ and try to build an equation for x. You'll have to use some trig, of course!

93. Compare $|w + z|^2$ to $(|w| + |z|)^2$. Remember, $|z|^2 = z \cdot \bar{z}$.

94. What is $\frac{1}{x+yi}$?

95. Let the first four lines meet at the origin.

96. Let the vertices of the triangle be the cube roots of unity. Start with PA^2, PB^2, and PC^2. Letting p correspond to P, write the sum of these squares in terms of p and $\bar{p}$.

97. Start with an easier problem; let $\mathbf{u} = \mathbf{i}$ and $\mathbf{v} = \mathbf{j}$.

98. Think about the problem geometrically. In each part, consider what the given conditions tell us about the possible locations of the head of $\mathbf{b}$.

99. Sines and cosines, sines and cosines.

100. Find a vector in terms of $\vec{B}$, $\vec{C}$, and $\vec{D}$ that is in the same direction as the altitude through A.

101. Build an equilateral triangle.

102. How is a real number related to its conjugate?

103. Can you pair up the angles in the numerator in a convenient way?

104. Try the Angle Bisector Theorem.

105. How can you relate $A_1\vec{A}_2$ to $B_1\vec{B}_2$?

106. Let $WXYZ$ be a square. What is the length of the projection onto $\vec{XY}$ of the vector from a point on side $\overline{WX}$ to a point on $\overline{YZ}$?

107. How would you parameterize the whole circle? What restrictions do you need in order to omit the part of the circle that you don't want?

108. You could get a closer look at this with a telescope.

109. What is $|z_1|$?

110. You have two right triangles (BAM and BDM) with related angles, and you have some side lengths. Trig away! Try to build an equation for θ.

111. The less fun way: Let c_1 and c_2 be the two possibilities for c. Find expressions for $c_1 + c_2$ and c_1c_2 in terms of a and b. What quadratic has c_1 and c_2 as roots?

112. Find $\det(\mathbf{A})$ in terms of v_1, v_2, θ, and $\|\mathbf{w}\| / \|\mathbf{v}\|$. When is this determinant negative?

113. What does $z^{28} - z^8 = 1$ tell us about the relationship between z^{28} and z^8 in the complex plane?

114. Suppose an octahedron has edge length x. What are the possible lengths of a segment connecting two vertices of the octahedron in terms of x?

115. Write the system of equations using matrices and vectors. Do you recognize the matrix?

116. What is the inverse of $\mathbf{A}$?

117. When finding a parameterization for the red point, focus first on its position *relative to the center of* $\mathcal{B}$.

118. For the extra challenge, think of a pair of transformations such that the order in which you apply the transformations doesn't affect the final result if both are applied.

119. What is $\mathbf{A}^2 \begin{pmatrix} 1 \\ 3 \end{pmatrix}$?

120. Let P be the foot of the altitude from Z to $\overline{BC}$. Use the conditions that $\overleftrightarrow{ZP} \perp \overleftrightarrow{BC}$ and that P is on $\overleftrightarrow{BC}$ to write two equations. Solve these equations for p.

121. Let $z = a + bi$, and use the definition $|z| = \sqrt{a^2 + b^2}$. For another approach, what is the graph of the equation $|z - 1| = |z + 3|$ on the complex plane?

122. Can $\sin^5 \theta$ ever be larger than $\sin^2 \theta$?

123. All of the entries in the matrix are even.

124. What complex number has argument $\arctan \frac{1}{3}$?

125. Use the described property of the function to write an equation in terms of a, b, and z. Remember, this equation must hold for all z.

126. Algebra: Let $x = \cos 36°$ and $y = \cos 72° = \sin 18°$. Use trig identities to relate x and y.

127. Let A be the origin. What must be true about $(m - n)/(n' - b')$ if and only if $\overline{MN} \perp \overline{B'N'}$?

128. Eliminate the denominators!

129. Two 60° angles; it sure would be nice if there were a third.

130. Yes, the algebra gets a little nasty on this problem. Don't give up too soon, and stay organized.

131. Choose several values of x to come up with a guess for the period.

132. What is $(a + bi)(c + di)$?

133. Find an expression for AB/BC in terms of θ. The first hint should be helpful for this.

134. Let A be the foot of the altitude from $(4, -2)$ to the line. How is the vector from A to $(4, -2)$ related to the line?

135. Any three vectors in $\mathbb{R}^2$ are linearly dependent.

136. What is the determinant of $\mathbf{A}^n$?

137. You can't find the actual solutions without a calculator, but you can find *sums* of pairs of the solutions.

138. If you know what $\text{proj}_{\mathbf{v}}\mathbf{w}$ is, can you figure out what $\mathbf{w}$ is?

139. This problem would be a whole lot easier if v or w were simply 1.

140. The vector $\vec{DC} \times \vec{DB}$ is normal to what plane?

141. Show that each root of $x^n - 1$ is a root of exactly one $\Phi_{d_i}(x)$, and vice versa.

142. Recall that $x^{n-1} + x^{n-2} + \cdots + x + 1 = (x - \omega)(x - \omega^2) \cdots (x - \omega^{n-1})$, where $\omega = e^{2\pi i/n}$.

143. Let z be a complex number (point in the complex plane), and let $f(z)$ be its image. How can we write the distance between these two points as an algebraic expression?

144. What is the signed area of the parallelogram with $\mathbf{Ti}$ and $\mathbf{Tj}$ as sides?

145. Make a clever substitution to take care of the 4^n term.

146. Find $g^{-1}(x)$ in terms of a, b, c, and d.

147. What identity do you know that involves tangent and secant?

148. Find a parameterization for the path of the *center* of $\mathcal{B}$.

149. Let B_1 and B_2 be the points on k_1 and k_2, respectively, such that B_1B_2 is the shortest distance between the two lines. How is $\overline{B_1B_2}$ related to the two lines?

150. Create equations involving p and $\overline{p}$ with the two conditions given for p: it is on the line through a and b, and the line through p and z is perpendicular to the line through a and b.

151. How can you interpret the problem geometrically?

152. Just for a moment, let $a = \alpha$, $b = \beta$, $c = a_n$, and $d = b_n$ in the right sides of the equations for a_{n+1} and b_{n+1}. Where have you seen expressions like that before when working with complex numbers?

153. What is $\tan(a + b + c)$?

154. $\pi/7$ is a fourteenth root of unity. Put on your complex number algebra hat.

155. The expression for $f(x)$ is almost a geometric series. What strategy do you use to prove the formula for the sum of a geometric series?

156. What must x be to make $\frac{x}{x-1}$ equal $\sec^2\theta$?

157. Geometry: Convert 18° to radians. What geometric figure gives you angles in radians with factors of 5 in the denominators?

158. Replace 3θ with another variable. What points on the unit circle are terminal points of angles that satisfy the resulting inequality?

159. Sines and cosines are usually easier to deal with than cotangents.

160. Let $\mathbf{A} = \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix}$, so that $\mathbf{v}$ and $\mathbf{w}$ are the columns of $\mathbf{A}$. Express the components of $\mathbf{w}$ in terms of $\|\mathbf{w}\|$, $\|\mathbf{v}\|$, and θ, where θ is the smallest counterclockwise rotation that must be applied to $\mathbf{v}$ to make the result have the same direction as $\mathbf{w}$.

161. Put the rectangle on the complex plane with one vertex at the origin and two sides on axes.

162. Can you write a trigonometric equation that has $\theta = 2\pi/7$, $\theta = 4\pi/7$, and $\theta = 6\pi/7$ as solutions?

163. You know how to find the cosine of the angle between two vectors.

164. What is true about the diagonals of a parallelogram?

165. The series is almost a geometric series. What strategy do you use to prove the formula for the sum of a geometric series?

166. Remember that y is constant. Can you write the expression in terms of only y?

167. Complex numbers are vectors.

168. Let $s = \sin\theta + \cos\theta$ and $p = (\sin\theta)(\cos\theta)$. Find two equations relating s and p, one with trig identities, and the other with the given equation.

169. Find two expressions for BD^2. Use them to find an expression of $a^2 - b^2 - c^2 + d^2$.

170. After moving all the terms to one side, what type of series do you have?

171. The cross-section given in the problem is not the only useful cross-section. Consider the cross-section that is perpendicular to the Earth's axis and includes Vancouver. Let A and B be the points on the circumference of this cross-section that divide the cross-section into "night" and "day." Where are points C and D in this cross-section?

172. Let P_k be ω^k and let z correspond to z. Express PP_k^2 in terms of z, $\bar{z}$, and ω.

173. The space diagonals of an octahedron are perpendicular.

174. Suppose ki is one of the imaginary roots. Let $x = ki$ in the equation. What must both the resulting real and imaginary parts equal?

175. Build triangles, and remember that equal angles have equal cosines.

176. The angle bisectors of a triangle meet at the incircle of the triangle. Build a right triangle with an angle that has measure $\frac{\angle A}{2}$.

177. Take it in steps. How would you parameterize the circle if it were centered at the origin?

178. The left side is the cube of a product. What product? And what sum can you write that product as?

179. When we say multiple, we mean multiple! As in, all of them. All 6 of them. Then, start from the relationships that involve the side you want (BE), and try to work as much as possible with sides you know and angles with equal sines.

180. You can find the angles of the hexagon. You then know how much you have to rotate each side in order to match the direction of an adjacent side. What tool works well with rotations?

181. If $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are linearly independent vectors in $\mathbb{R}^3$, then any vector in $\mathbb{R}^3$ can be written as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$.

182. Start with a polynomial that has all 12 twelfth roots of unity as roots.

183. You answer to first hint should give you expressions of the form $e^{i\alpha} - e^{-i\alpha}$. We'd rather deal with expressions of the form $1 - e^{i\alpha}$, since we've seen these sorts of expressions before. (Where?)

184. What is $f(t)$?

185. We have $f(f(x)) = x$ for *all* x. Choose a convenient x.

186. Start with two triangles, $\triangle ABD$ and $\triangle BCD$, as in the first part. Notice that these triangles share a side.

187. Draw a diagram that has a right triangle with one side of length $\sin\frac{6\pi}{7}$ and hypotenuse of length 1. Find an angle in the diagram that has cosine equal to $\sin\frac{6\pi}{7}$.

188. Show that each root of $\Phi_{2n}(x)$ is a root of $\Phi_n(-x)$, and vice versa. What are the negatives of the roots of $\Phi_n(-x)$?

189. Try using a few different identities to rewrite the each of the first three or four terms in the sum. Do any of your methods allow you to combine terms conveniently?

190. Consider the cubic polynomial with roots $e^{i\alpha}$, $e^{i\beta}$, and $e^{i\gamma}$. Expand it—what do you know about the resulting coefficients?

191. You probably don't need a hint to try complex numbers. But you might need a hint to deal with the expression you'll get for O_1O_2. Here are a couple: if $\omega = e^{\pi i/13}$, then what is ω^{13} and what power of ω equals $\overline{\omega^k}$?

192. What complex number has argument $\arctan 2$?

193. Buy low and sell high!!! How many times can you do so?

194. In what other area formula do we have the product of two side lengths? In what other triangle formula do we have the circumradius?

195. Try turning each product into a sum.

196. We have a triangle area formula that involves sine. We also found a vector formula that involves sine.

197. Let t be the desired length. You should be able to find all the dot products in your equation, or to express them in terms of t.

198. Can you find the area of the triangle with the three given points as vertices?

199. Let $\vec{AB} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, $\vec{AC} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$, and $\vec{AD} = d_1\mathbf{i} + d_2\mathbf{j} + d_3\mathbf{k}$. For each part, what determinant (in terms of the components of these vectors), equals the desired volume of the described parallelepiped?

200. Show that $h_2 = i(w - z)\cot\angle WOZ$.

201. As an Exercise, you should have learned how to deal with angle bisectors of a triangle with vectors.

202. What transformation has no final effect if applied three times?

203. $A + B + C = 180°$. Expressions with two variables are often easier to work with than expressions with three variables!

204. The professor entered t, and should have entered the expression you found in the first hint. And yet the professor got the right answer! So, what equation does that give you?

205. Find the angles at P. Then what equations can you write?

206. Experiment! Compute a few of the z_i. Notice anything interesting?

207. The expression $\sqrt{x^2 + (x-1)^2}$ looks like the distance formula.

208. What trig identity does the right side of the equation resemble?

209. That sum can be written in a much more concise form!

210. Show that $f(\mathbf{w}) = \text{proj}_{\mathbf{v}}\mathbf{w}$ is a linear function.

211. How can you relate sines and cosines to exponentials?

212. What identity does the expression $(x - y)/(1 + xy)$ resemble? What substitution might you make for x and y to recast this problem as a problem involving angles instead of one involving real numbers?

213. If $\det(\mathbf{A} - \lambda\mathbf{I}) \neq 0$, then what vectors $\mathbf{v}$ satisfy the equation $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$?

214. Find a convenient way to pair up the terms.

215. What choice of origin allows you to take advantage of all those right angles?

216. Let $b_n = a_n/2^n$. Write the recursion in terms of b_{n+1} and b_n. What substitution for b_n does the resulting recursion suggest?

217. How are the $x_k + y_k i$ related to the tenth roots of unity?

218. Focus on the two circles that produce Z_3. Use similar triangles to express z_3 in terms of the radii of the circles and the complex numbers corresponding the centers of the circles.

219. Why must $\angle SOC = 66.5°$ on the longest day of the year in Vancouver? (If you can't figure this out, assume it is true and try to finish the problem from here.)

220. Any point on an angle bisector of an angle is equidistant from the sides of the angle.

221. Express $\mathbf{A}^2$ in terms of $\mathbf{A}$, $\mathbf{I}$, and entries in $\mathbf{A}$.

222. Forget about matrices for a minute. What pairs of integers (a, b) satisfy $ab = a + b$? Try to find a solution that doesn't involve division; there's no "division" in matrices!

223. Product of terms. Double angles. Which identity should you try to use?

224. How are $\angle AOB$ and $\angle ACB$ related in an acute triangle ABC? A right triangle with $\angle C = 90°$? An obtuse triangle with $\angle C > 90°$?

225. Here are two first steps you might consider for turning the equation into something simpler you can handle: (1) Write the equation in terms of sine and cosine, (2) View the equation as a quadratic in $\tan x$. Again, remember that you don't have to find x!

226. The a^2+b^2 suggests complex numbers. Let $z = a+bi$. Can you write an equation in terms of z that is equivalent to the given system?

227. How can you factor $x^n - 1$ as the product of n polynomials? How can you factor $x^n - 1$ as the product of 2 polynomials?

228. Let $PB = x$. Can you find PA' in terms of x?

229. Two angles that are inscribed in the same arc of a circle are equal.

230. What is $\mathbf{A}^3$?

231. How is $\mathbf{v} - \vec{F}$ related to the plane?

232. Let the hexagon be $ABCDEF$. Write equations involving vectors to the vertices. Can you group the equations conveniently?

233. The more fun way: What cubic has roots a, b, and c? What simple cubic has roots that are the vertices of an equilateral triangle?

234. Find $\angle BOC$ first.

235. After solving the problem in which the hole is drilled along the z-axis, how can you quickly change your result to account for a hole being along the x-axis?

236. Where have you seen an expression of the form $(x - \omega)(x - \omega^2)(x - \omega^3)\cdots$?

237. Draw a picture.

238. Find other angles in terms of θ. Did you discover anything interesting?

239. After you have a guess that the period is p, use the unit circle to show that we must always have $f(x+p) = f(x)$. (Focus on sine and cosine separately, and don't forget to show that no smaller period is possible.)

240. Solve the first equation for y, the second for z, and the third for x. What trig identity resembles the resulting equations?

241. There are three pairs of angles in the problem with equal sines. Try applying the Law of Sines to multiple triangles, focusing on sides you know and these pairs of angles.

242. Find the volume of the parallelepiped with $\vec{AB}$, $\vec{AC}$, and $\vec{AD}$ as edges. Try to relate the desired volumes to this volume.

243. You'll need to be able to express the center of an equilateral triangle on the complex plane in terms of its vertices. The midpoint of a segment is the average of its endpoints. What do you think the center of an equilateral triangle is?

244. First, consider the problem in which the hole is drilled along the z-axis.

245. Notice anything interesting about that particular matrix?

246. The equation isn't in the form of an ellipse with a horizontal major axis. Maybe it's the result of rotating an ellipse with horizontal major axis about the origin?

247. Try complex numbers. Show that $(h_2 - h_1)/(g_2 - g_1)$ is imaginary.

248. The Law of Cosines is not the only way to introduce cosine; what's $\sin 2\theta$?

249. Focus on the four key points: the two ends of the graph and the two points where the graph changes direction. Where are these points on the new graph?

250. Because you're looking back here, you've discovered that this problem is a lot harder than it looks! Find lots of angles, and keep your eye out for supplementary angles. What trig law can we sometimes use in problems with convenient supplementary angles?

251. Let A be the origin. Write an equation for the given perpendicularity.

252. The seventh roots of unity are the roots of what polynomial?

253. Note that x and y can each be any real number. So can the tangent of an angle.

254. If $x^2 + x + 1 = 0$, then how can you write $(x + 1)^{2n}$ as a power of x?

255. The expression $\dfrac{\mathbf{v} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a}$ should be familiar—geometrically speaking, how is this expression related to $\mathbf{a}$ and $\mathbf{v}$?

256. There's a reason this part comes after part (a)!

257. The equation is just a quadratic; you can find z!

258. There are a lot of right angles at D.

259. Many 3D triangles are 2D problems in disguise. What does the problem look like if you unroll the curved surface of the cone?

260. Break the quadrilateral into two triangles, $\triangle ABD$ and $\triangle BCD$. How are $\sin A$ and $\sin C$ related? And how can you get these from $\cos A$?

261. Sine and cosine are often easier to deal with than tangent.

262. Find two quadratics whose product is the given quartic.

263. Use wishful thinking. Can you change a single term in the polynomial into something else to make the polynomial into one you can handle? What's the difference between the polynomial you are given and the one you can handle?

264. For each of the first four lines, write an equation that means that the origin is on the line. Be careful to write them so that you can use convenient symmetry in your equations!

265. Creatively add segments to $\triangle ABC$ to create an angle equal to $\angle B - \angle C$.

266. Connect all three points to the center.

267. Two of the terms in the equation are reciprocals. What is the smallest possible value of the sum of a positive number and its reciprocal?

268. Write the equation for the line in terms of w, then let $w = a + bi$.

269. Let A be the origin and find every point you can in terms of b and c.

270. Focus on linearly dependent vectors first, since that gives you an equation to work with: $a\mathbf{p} + b\mathbf{q} = \mathbf{0}$ for some constants a and b that are not both 0.

271. Try some specific examples, such as $f(x) = x^3$ or $f(x) = x + 3$.

272. The left sides look like the Binomial Theorem.

273. Cosines and vectors. What tool should you consider?

274. Why must there be scalars p_1, q_1, p_2, q_2 such that $\mathbf{v} = p_1\mathbf{a} + q_1\mathbf{b}$ and $\mathbf{w} = p_2\mathbf{a} + q_2\mathbf{b}$?

275. Let X be the intersection of the lines containing the altitudes from C and E. Suppose $\vec{AX} \cdot \vec{FB} = k$. What do we wish to prove k equals?

276. Suppose the difference in arguments between v and w is α. Show that $|v + w| = |1 + e^{i\alpha}|$.

277. How are ab and $a + b$ related to $(a - 1)(b - 1)$?

278. Area.

279. Draw altitude $\overline{BY}$. What angle of the triangle equals $\angle AHY$?

280. We have two cosines; try building a right triangle.

281. Compare this equation to the condition we proved in the text under which a, b, and c are the vertices of an equilateral triangle in the complex plane. What must be true about ω and ζ if both equations are true?

282. Use trig identities to express the right side as the tangent of an angle.

283. All the variable terms are powers of z^3. If you let $t = z^3$, do you get a polynomial you recognize?

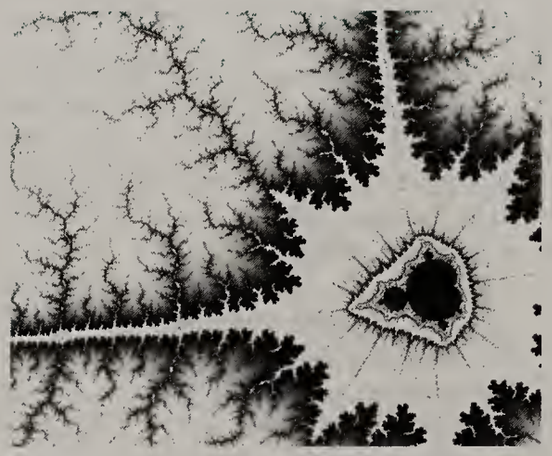
284. Choose your origin wisely. Specifically, choose it in a way that wipes out a lot of the terms in the expression you must show equals 0 in order to show that z_1, z_2, and z_3 are collinear.

285. It's the last problem of the book—you're on your own!

286. Your expression for $[ABCD]$ from the first hint should look vaguely similar to your expression for $a^2 - b^2 - c^2 + d^2$. One should have sines where the other has cosines. One way to make sines and cosines go away is $\sin^2\theta + \cos^2\theta = 1$.

287. Compare the expression the Shoelace Theorem gives for $n = k$ to the result when $n = k + 1$.

288. What is $|w + z|^2 - (|w| + |z|)^2$?



Index

α, 29
$\setminus$, 6
β, 29
$\cap$, 6
$\cup$, 5
∞, 2
ω, 29
$\perp$, 310
ϕ, 179
ρ, 179
θ, 29
30-60-90 triangle, 32
45-45-90 triangle, 32

Agnesi, Maria Gaetana, 191
Airy, Sir George, 180
AMC, vii
American Mathematics Competitions, *see* AMC
American Regions Math League, *see* ARML
amplitude, 56
angle
    dihedral, 155
    inscribed, 141
Angle Bisector Theorem, 155
angle difference identity, 91–99
angle of depression, 129, 130
angle of elevation, 129, 130
angle sum identity, 91–99
angular coordinate, 167
arccos, 68
arccot, 68
arccsc, 68
arcsec, 68
arcsin, 68
arctan, 68
argument, 220
ARML, vii
Art of Problem Solving, v, 512
associative property
    complex numbers, 195
asymptote, 50

bilinear, 308, 383
box product, 451
Brahmagupta's formula, 157
Bretschneider's formula, 157

Cartesian space, 177
Cauchy-Schwarz Inequality, 311
centroid, 272, 470
    of a tetrahedron, 475
    vector representation, 470
Chebyshev polynomial, 224
circle
    complex number representation, 201, 274
circle of Apollonius, 275
circular reasoning, 231
circumcircle, 144, 148
circumradius, 144, 148, 474
cis, 220
cofactor, 405

collinearity, 266, 270
  complex numbers and, 266, 270
column vector, 343, 409
commutative property
  complex numbers, 195
complex number, 192–297
  argument, 220
  arithmetic, 192–197
  associative property, 195
  commutative property, 195
  complex plane, 198–202
  conjugate, 195
  distributive property, 195
  equilateral triangle, 280
  exponential form, 225–233
  geometry, 256–297
  graphing, 198–202
  imaginary part, 192, 202–208
  magnitude, 200, 220
  multiplicative inverse, 196
  norm, 200
  parallelogram law, 217
  polar form, 219–225
  real part, 192, 202–208
  rectangular form, 220
  root of unity, 233–249
  square root, 193, 206
complex number geometry, 256–297
  circle, 201, 274
  collinearity, 266, 270
  midpoint, 264
  parallel lines, 267, 270
  perpendicular lines, 268, 270
  regular polygon, 278–286
complex plane, 198–202
component, 298
composition (of functions), 16–19
conjugate, 195
conjugate pair, 195
converse, 141
conversion factor, 41
coordinates
  cylindrical, 178–187
  polar, 167–176
  rectangular, 178
  spherical, 179, 187
cos, 30
$\cos^{-1}$, 68
cosecant, 30
  graph of, 52
cosine, 30
  angle difference identity, 95
  angle sum identity, 95
  difference-to-product identity, 109
  double angle identity, 101
  graph of, 48
  half angle identity, 104
  in terms of e, 231
  period of, 47
  sum-to-product identity, 109
  triple angle identity, 106
cot, 30
cotangent, 30
Cramer's rule, 357, 397
cross product, 442–449
csc, 30
cyclic quadrilateral, 157
cycloid, 165
cyclotomic polynomial, 253
cylindrical coordinates, 178–187

de Moivre's Theorem, 223, 227
degree, 212
determinant, 357
  geometric interpretation (2×2), 364–369
  geometric interpretation (3×3), 450–454
  of a 2×2 matrix, 355–369
  of a 3×3 matrix, 395–408, 450–454
diagonal matrix, 348
difference-to-product identity, 106–110
dihedral angle, 155
dilation, 257
  center of, 257
  scale factor, 257
dimensionless unit, 40
direction vector, 316, 423
distributive property
  complex numbers, 195
domain, 1
dot product
  in three dimensions, 382–384
  in two dimensions, 305–312
double angle identity, 100–106
double root, 217
dummy variable, 1

e, 227–228
Edwards, Miles Dillon, 484
eigenvalue, 419
eigenvector, 419
ellipse, 175
equation
  parametric, 158–167
equator, 180

equilateral triangle
    complex numbers and, 280
Euler line, 471, 472
even function, 50
excircle, 474
expansion by minors, 404
exponential form, 225–233
Extended Law of Sines, 144
extraneous solution, 85

f^{-1} (functions), 20
Factor Theorem, 209
Feuerbach's Theorem, 474
fractals, xiv
frequency, 61, 111
function, 1–28
    composition, 16–19
    domain, 1
    even, 50
    graphing, 9–16
    inverse, 19–24
    invertible, 22
    linear, 308, 311
    odd, 50
    periodic, 46
    range, 1
    real, 2
    real-valued, 2
    transformation, 12–16
    trigonometric, 29
Fundamental Theorem of Algebra, 209

generalization, 138
generated, 426, 427
geometric transformation, 256
    complex numbers and, 256–265
geometry
    complex number, 256–297
great circle, 180
Greek letter, 29
guitar, 111

half angle identity, 100–106
Harvard-MIT Math Tournament, *see* HMMT
head, of a vector, 303
head-to-tail addition, 303
Heron's formula, 152
    proof with complex numbers, 485
HMMT, viii
horizontal line test, 23
hyperbolic cosine, 252
hyperbolic sine, 252
hyperbolic tangent, 252

i, 192
identity (trigonometric), 38, 83–127
    angle difference, 91–99
    angle sum, 91–99
    difference-to-product, 106–110
    double angle, 100–106
    half angle, 100–106
    product-to-difference, 106–110
    product-to-sum, 106–110
    sum-to-product, 106–110
    triple angle, 106
identity matrix, 350, 393
    2×2, 350
    3×3, 393
if and only if, 204
i^i, 255
Im, 202
imaginary axis, 198
imaginary number, 192
imaginary part, 192, 202–208
incircle, 148, 474
inner product, 306
inradius, 148, 474
inscribed angle, 141
instrument tuning, 111
intercept, 9
intersection, of intervals, 6
interval
    intersection, 6
    union, 6
interval notation, 2
inverse cosine, 68
inverse function, 19–24
    graph of, 23
inverse matrix (2×2), 369–375
inverse matrix (3×3), 454–457
inverse sine, 68
inverse tangent, 68
inverse trigonometric function, 64–76
invertible, 372
irreducible polynomial, 242
isometry, 257

Jacobi Identity, 449

kite, 469

Lagrange Identity, 449
latitude, 180
lattice point, 369
Law of Cosines, 136–141

Law of Sines, 141–148
Law of Tangents, 152
length
    of a vector, 381
line
    definition of, two dimensions, 316
    distance from a point to, 330
linear combination, 319, 421
linear function, 308, 311
linearly dependent, 322, 441
linearly independent, 322, 441
links, vi
logarithm
    base e, 255
longitude, 180

magnitude, 220
    of a complex number, 200
    of a vector, 381
Mandelbrot Competition, vii
Mandelbrot set, xv
matrices, 335
matrix
    2×2, 335–379
    3×3, 384–408
    addition, 338
    cofactor, 405
    determinant, 355–369, 395–408, 450–454
    diagonal, 348
    dimensions, 409
    eigenvalue, 419
    eigenvector, 419
    identity, 350, 393
    inverse (2×2), 369–375
    inverse (3×3), 454–457
    multiplication, 342
    nonsingular, 362
    permutation, 394
    rotation, 352
    singular, 362
    skew-symmetric, 419
    square, 414
    transpose, 378, 399
    zero, 336
median, 272
    of a tetrahedron, 475
midpoint
    in the complex plane, 264
minor, 404
Monge's Theorem, 289
monic, 212
multiplicative inverse, 196

New York State Math League, viii
nine-point circle, 471
    complex numbers, 479
nonreal, 192
nonsingular matrix, 362, 397
norm
    in three dimensions, 381
    in two dimensions, 300
    of a complex number, 200
normal, 323, 438
number
    complex, 192
    imaginary, 192
    nonreal, 192
    real, 192
NYSML, viii

odd function, 50
orientation, 256
origin, 198
orthocenter, 471
    complex numbers, 479
    vector representation, 471
orthogonal, 310, 384
Outer Napoleon Triangle, 288

parallel lines
    complex numbers and, 267, 270
parallel planes, 427
parallelepiped, 450
parallelogram
    area of, 449
parallelogram law, 217
parameter, 158
parameterization, 163
parameterize an equation, 164
parametric equation, 158–167
period, 46
periodic function, 46
permutation matrix, 394
perpendicular lines
    complex numbers and, 268, 270
phase shift, 58
Pigeonhole Principle, 119
pirate, 263
pitch, 111
plane, 427
    normal, 438
    parallel, 427
    through the origin, 426
polar coordinates, 167–176
polar form, 219–225

polynomial
    complex roots of, 208–214
    cyclotomic, 253
    degree, 212
    Factor Theorem, 209
    irreducible, 242
    monic, 212
    Rational Root Theorem, 209
prime meridian, 180
primitive root of unity, 238
problem solving, iii
product-to-difference identity, 106–110
product-to-sum identity, 106–110
proj, 323, 328
projection, 323–331
prosthaphaeresis, 124
Ptolemy's Inequality, 287
Pythagorean Theorem, 39

quadrant, 35

$\mathbb{R}$, 2
$\mathbb{R}\backslash$, 6
radial coordinate, 167
radian, 40–44
range, 1
Rational Root Theorem, 209
Re, 202
real axis, 198
real function, 2
real number, 192
real part, 192, 202–208
rectangular coordinates
    three dimensions, 178
    two dimensions, 167
rectangular form, 220
reflection, 257
regular polygon
    complex numbers and, 278–286
resources, v
right triangle trigonometry, 35, 128–136
right-hand rule, 445
root of unity, 233–249
    primitive, 238
rotation, 256
    angle, 256
    center, 256
    orientation, 256
rotation matrix, 352
row vector, 343, 409
Rule of Sarrus, 397

scalar, 298
scalar triple product, 451
scale factor, 257
sec, 30
secant, 30
semiperimeter, 148
Shoelace Theorem, 379
signed area, 367
signed volume, 451
Simson Line, 291
sin, 30
$\sin^{-1}$, 68
sine, 30
    angle difference identity, 95
    angle sum identity, 95
    difference-to-product identity, 109
    double angle identity, 101
    graph of, 47
    half angle identity, 104
    in terms of e, 231
    sum-to-product identity, 109
    triple angle identity, 106
sine curve, 48
singular matrix, 362, 397
sinusoid, 48
sinusoidal variation, 61
skew-symmetric matrix, 419
slope, 30, 131
SOHCAHTOA, 128
sound, 111
sound waves, 111
spherical coordinates, 179, 187
square matrix, 414
Stewart's Theorem, 151
sum-to-product identity, 106–110

tail, of a vector, 303
tan, 30
$\tan^{-1}$, 68
tangent, 30
    and slope, 131
    angle difference identity, 95
    angle sum identity, 95
    double angle identity, 101
    graph of, 51
    half angle identity, 104
Tchebychev polynomial, 224
telescoping series, 116
terminal point, 30
tetrahedron
    centroid, 475
    median, 475
three-dimensional space, 177

transformation, 256
translation, 256
transpose, 378, 399, 419
    of a 2×2 matrix, 378
    of a 3×3 matrix, 399
    of an $r \times c$ matrix, 419
triangle area, 133
Triangle Inequality, 275, 277
trig, 30
trigonometric function, 29
    graph of, 45–64
    in a right triangle, 35
    inverse of, 64–76
trigonometric identity, 38, 83–127
trigonometry, 29–157
    Extended Law of Sines, 144
    Law of Cosines, 136–141
    Law of Sines, 141–148
    Law of Tangents, 152
    right triangle, 128–136
triple angle identity, 106

union, of intervals, 6
unit circle, 30
unit vector, 299, 381
upper triangular matrix, 408, 457
USA Mathematical Talent Search, *see* USAMTS
USAMTS, vii

variable
    dummy, 1
variation
    sinusoidal, 61
vector, 298
    centroid of a tetrahedron, 475
    centroid of a triangle, 470
    column, 343
    component, 298
    cross product, 442–449
    direction, 302, 381
    dot product, 305–312, 382–384
    head, 303
    head-to-tail addition, 303
    in three dimensions, 380–384
    in two dimensions, 298–334
    inner product, 306
    length, 300, 381
    linear dependence, 313–322
    magnitude, 300, 381
    norm, 300, 381
    normal, 323, 438
    orthocenter, 471
    projection, 323–331
    row, 343
    tail, 303
    unit, 299, 381
    zero, 299
vertical line test, 12
Vieta's formulas, 246

witch of Agnesi, 191

x-intercept, 9
xy-plane, 177
xz-plane, 177

y-intercept, 9
yz-plane, 177

$\bar{z}$, 195
zero matrix, 336
zero vector, 299
zero-product property, 210

www.artofproblemsolving.com

The Art of Problem Solving (AoPS) is:

• Books

For over 16 years, the classic *Art of Problem Solving* books have been used by students as a resource for the American Mathematics Competitions and other national and local math events.

> *Every school should have this in their math library.*
> – Paul Zeitz, past coach of the U.S. International Mathematical Olympiad team

The Art of Problem Solving Introduction and Intermediate texts, together with our new *Precalculus* and *Calculus* texts, form a complete curriculum for outstanding math students in grades 6-12.

> *The new book [Introduction to Counting & Probability] is great. I have started to use it in my classes on a regular basis. I can see the improvement in my kids over just a short period.*
> – Jeff Boyd, 4-time MATHCOUNTS National Competition winning coach

• Classes

The Art of Problem Solving offers online classes on topics such as number theory, counting, geometry, algebra, and more at beginning, intermediate, and Olympiad levels.

> *All the children were very engaged. It's the best use of technology I have ever seen.*
> – Mary Fay-Zenk, coach of National Champion California MATHCOUNTS teams

• Forum

As of November 2009, the Art of Problem Solving Forum has over 71,000 members who have posted over 1,600,000 messages on our discussion board. Members can also participate in any of our free "Math Jams."

> *I'd just like to thank the coordinators of this site for taking the time to set it up... I think this is a great site, and I bet just about anyone else here would say the same...*
> – AoPS Community Member

• Resources

We have links to summer programs, book resources, problem sources, national and local competitions, scholarship listings, a math wiki, and a LaTeX tutorial.

> *I'd like to commend you on your wonderful site. It's informative, welcoming, and supportive of the math community. I wish it had been around when I was growing up.*
> – AoPS Community Member

• . . . and more!

Membership is **FREE**! Come join the Art of Problem Solving community today!

the Art of Problem Solving
Precalculus

Precalculus is part of the acclaimed Art of Problem Solving curriculum designed to challenge high-performing middle and high school students. **Precalculus** covers trigonometry, complex numbers, vectors, and matrices. It includes nearly 1000 problems, ranging from routine exercises to extremely challenging problems drawn from major mathematics competitions such as the American Invitational Mathematics Exam and the USA Mathematical Olympiad. Almost half of the problems have full, detailed solutions in the text, and the rest have full solutions in the accompanying Solutions Manual.

"Every time one of your books arrives on my doorstep I couldn't be more impressed with what you're doing... For the right kind of student—and I was one myself, I just wish these books had been around when I was a kid—your books must be like a prayer answered. And as I've always said to every math teacher I meet, this is the way to teach mathematics to everyone."
Steve Olson, author of *Count Down*

Richard Rusczyk is the founder of www.artofproblemsolving.com. He is co-author of *the Art of Problem Solving, Volumes 1 and 2,* and *Intermediate Algebra* and author of *Introduction to Algebra* and *Introduction to Geometry*. He was a national MATHCOUNTS participant, a three-time participant in the Math Olympiad Summer Program, a perfect scorer on the AIME, and a USA Math Olympiad Winner.

ISBN 978-1-934124-16-1
90000 >
9 781934 124161